Technology for the United States Navy and Marine Corps, 2000-2035

Becoming a 21st-Century Force

VOLUME 9 Modeling and Simulation

NAVAL STUDIES BOARD
NATIONAL RESEARCH COUNCIL

"Technology for the U.S Navy and Marine Corps - 2000 - 2035 Becoming a 21st Century Force"

PLEASE CHECK THE APPROPRIATE BLOCK BELOW:

-AQ# U99-05-0372 Vol 9

☐ ____________ copies are being forwarded. Indicate whether Statement A, B, C, D, E, F, or X applies.

☒ DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED — all 9 volumes

☐ DISTRIBUTION STATEMENT B:
DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES ONLY; (Indicate Reason and Date). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office).

☐ DISTRIBUTION STATEMENT C:
DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND THEIR CONTRACTORS; (Indicate Reason and Date). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office).

☐ DISTRIBUTION STATEMENT D:
DISTRIBUTION AUTHORIZED TO DoD AND U.S. DoD CONTRACTORS ONLY; (Indicate Reason and Date). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office).

☐ DISTRIBUTION STATEMENT E:
DISTRIBUTION AUTHORIZED TO DoD COMPONENTS ONLY; (Indicate Reason and Date). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office).

☐ DISTRIBUTION STATEMENT F:
FURTHER DISSEMINATION ONLY AS DIRECTED BY (Indicate Controlling DoD Office and Date) or HIGHER DoD AUTHORITY.

☐ DISTRIBUTION STATEMENT X:
DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND PRIVATE INDIVIDUALS OR ENTERPRISES ELIGIBLE TO OBTAIN EXPORT-CONTROLLED TECHNICAL DATA IN ACCORDANCE WITH DoD DIRECTIVE 5230.25. WITHHOLDING OF UNCLASSIFIED TECHNICAL DATA FROM PUBLIC DISCLOSURE. 6 Nov 1984 (Indicate date of determination). CONTROLLING DoD OFFICE IS (Indicate Controlling DoD Office).

☐ This document was previously forwarded to DTIC on ____________ (date) and the AD number is ________________.

☐ In accordance with provisions of DoD instructions, the document requested is not supplied because:

☐ It will be published at a later date. (Enter approximate date, if known).

☐ Other. (Give Reason)

DoD Directive 5230.24, "Distribution Statements on Technical Documents," 18 Mar 87, contains seven distribution statements, as described briefly above. Technical Documents must be assigned distribution statements.

MARY GERDEL
Print or Type Name

Technology for the United States Navy and Marine Corps, 2000-2035

Becoming a 21st-Century Force

VOLUME 9 Modeling and Simulation

Panel on Modeling and Simulation
Committee on Technology for Future Naval Forces
Naval Studies Board
Commission on Physical Sciences, Mathematics, and Applications
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1997

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce Alberts and Dr. William A. Wulf are chairman and vice chairman, respectively, of the National Research Council.

This work was performed under Department of the Navy Contract N00014-96-D-0169/0001 issued by the Office of Naval Research under contract authority NR 201-124. However, the content does not necessarily reflect the position or the policy of the Department of the Navy or the government, and no official endorsement should be inferred.

Copies available from:

Naval Studies Board
National Research Council
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Printed in the United States of America

PANEL ON MODELING AND SIMULATION

GEORGE F. CARRIER, Harvard University, *Chair*
PAUL K. DAVIS, RAND and the RAND Graduate School, *Vice Chair*
DONALD K. BLUMENTHAL, Gualala, California
RICHARD BRONOWITZ, Center for Naval Analyses
JOHN C. DOYLE, California Institute of Technology
DONALD P. GAVER, Naval Postgraduate School
DON E. HIHN, Charleston, South Carolina
RICHARD J. IVANETICH, Institute for Defense Analyses
JOHN P. LEHOCZKY, Carnegie Mellon University
DAVID L. McDOWELL, Georgia Institute of Technology
DUNCAN C. MILLER, Massachusetts Institute of Technology
DAVID R. OLIVER, Northrop Grumman Corporation
GABRIEL ROBINS, University of Virginia
BERNARD P. ZEIGLER, University of Arizona

Invited Participant
BEN P. WISE, Science Applications International Corporation

Navy Liaison Representatives
CDR THOMAS COSGROVE, USN, Office of the Chief of Naval Operations, N858D
CAPT JAY KISTLER, USN, Office of the Chief of Naval Operations, N6M

Consultants
LEE M. HUNT
SIDNEY G. REED, JR.
JAMES G. WILSON

Staff
RONALD D. TAYLOR, Director, Naval Studies Board
PETER W. ROONEY, Program Officer
SUSAN G. CAMPBELL, Administrative Assistant
MARY G. GORDON, Information Officer
CHRISTOPHER A. HANNA, Project Assistant

COMMITTEE ON TECHNOLOGY FOR FUTURE NAVAL FORCES

DAVID R. HEEBNER, Science Applications International Corporation (retired), *Study Director*
ALBERT J. BACIOCCO, JR., The Baciocco Group, Inc.
ALAN BERMAN, Applied Research Laboratory, Pennsylvania State University
NORMAN E. BETAQUE, Logistics Management Institute
GERALD A. CANN, Raytheon Company
GEORGE F. CARRIER, Harvard University
SEYMOUR J. DEITCHMAN, Institute for Defense Analyses (retired)
ALEXANDER FLAX, Potomac, Maryland
WILLIAM J. MORAN, Redwood City, California
ROBERT J. MURRAY, Center for Naval Analyses
ROBERT B. OAKLEY, National Defense University
JOSEPH B. REAGAN, Saratoga, California
VINCENT VITTO, Lincoln Laboratory, Massachusetts Institute of Technology

Navy Liaison Representatives
RADM JOHN W. CRAINE, JR., USN, Office of the Chief of Naval Operations, N81 (as of July 4, 1996)
VADM THOMAS B. FARGO, USN, Office of the Chief of Naval Operations, N81 (through July 3, 1996)
RADM RICHARD A. RIDDELL, USN, Office of the Chief of Naval Operations, N91
CDR DOUGLASS BIESEL, USN, Office of the Chief of Naval Operations, N812C1
PAUL G. BLATCH, Office of the Chief of Naval Operations, N911E

Marine Corps Liaison Representative
LtGen PAUL K. VAN RIPER, USMC, Marine Corps Combat Development Command

Consultants
LEE M. HUNT
SIDNEY G. REED, JR.
JAMES G. WILSON

Staff
RONALD D. TAYLOR, Director, Naval Studies Board
PETER W. ROONEY, Program Officer
SUSAN G. CAMPBELL, Administrative Assistant
MARY G. GORDON, Information Officer
CHRISTOPHER A. HANNA, Project Assistant

NAVAL STUDIES BOARD

DAVID R. HEEBNER, Science Applications International Corporation (retired), *Chair*
GEORGE M. WHITESIDES, Harvard University, *Vice Chair*
ALBERT J. BACIOCCO, JR., The Baciocco Group, Inc.
ALAN BERMAN, Center for Naval Analyses
NORMAN E. BETAQUE, Logistics Management Institute
NORVAL L. BROOME, Mitre Corporation
GERALD A. CANN, Raytheon Company
SEYMOUR J. DEITCHMAN, Institute for Defense Analyses (retired), *Special Advisor*
ANTHONY J. DeMARIA, DeMaria ElectroOptics Systems, Inc.
JOHN F. EGAN, Lockheed Martin Corporation
ROBERT HUMMEL, Courant Institute of Mathematical Sciences, New York University
DAVID W. McCALL, Far Hills, New Jersey
ROBERT J. MURRAY, Center for Naval Analyses
ROBERT B. OAKLEY, National Defense University
WILLIAM J. PHILLIPS, Northstar Associates, Inc.
MARA G. PRENTISS, Jefferson Laboratory, Harvard University
HERBERT RABIN, University of Maryland
JULIE JCH RYAN, Booz, Allen and Hamilton
HARRISON SHULL, Monterey, California
KEITH A. SMITH, Vienna, Virginia
ROBERT C. SPINDEL, Applied Physics Laboratory, University of Washington
DAVID L. STANFORD, Science Applications International Corporation
H. GREGORY TORNATORE, Applied Physics Laboratory, Johns Hopkins University
J. PACE VanDEVENDER, Prosperity Institute
VINCENT VITTO, Lincoln Laboratory, Massachusetts Institute of Technology
BRUCE WALD, Arlington Education Consultants

Navy Liaison Representatives
RADM JOHN W. CRAINE, JR., USN, Office of the Chief of Naval Operations, N81 (as of July 4, 1996)
VADM THOMAS B. FARGO, USN, Office of the Chief of Naval Operations, N81 (through July 3, 1996)
RADM RICHARD A. RIDDELL, USN, Office of the Chief of Naval Operations, N91
RONALD N. KOSTOFF, Office of Naval Research

Marine Corps Liaison Representative
LtGen PAUL K. VAN RIPER, USMC, Marine Corps Combat Development Command

RONALD D. TAYLOR, Director
PETER W. ROONEY, Program Officer
SUSAN G. CAMPBELL, Administrative Assistant
MARY G. GORDON, Information Officer
CHRISTOPHER A. HANNA, Project Assistant

COMMISSION ON PHYSICAL SCIENCES, MATHEMATICS, AND APPLICATIONS

Preface

This report is part of the nine-volume series entitled *Technology for the United States Navy and Marine Corps, 2000-2035: Becoming a 21st-Century Force.* The series is the product of an 18-month study requested by the Chief of Naval Operations, who, in a memorandum on November 28, 1995, asked the National Research Council to initiate through its Naval Studies Board a thorough examination of the impact of advancing technology on the form and capability of the naval forces to the year 2035. To carry out this study, eight technical panels were organized under the committee on Technology for Future Naval Forces to examine all of the specific technical areas called out in the terms of reference.

The study's terms of reference (Appendix A) asked for an identification of "present and emerging technologies that relate to the full breadth of Navy and Marine Corps mission capabilities," with specific attention to "(1) information warfare, electronic warfare, and the use of surveillance assets; (2) mine warfare and submarine warfare; (3) Navy and Marine Corps weaponry in the context of effectiveness on target; [and] (4) issues in caring for and maximizing effectiveness of Navy and Marine Corps human resources." The terms of reference went on to identify 10 technical areas for special attention. One involved modeling and simulation (M&S): "The naval service is increasingly dependent upon modeling and simulation. The study should review the overall architecture of models and simulation in the DoD (DoN, JCS, and OSD), the ability of the models to represent real world situations, and their merits as tools upon which to make technical and force composition decisions."

It was against this background that the Panel on Modeling and Simulation was constituted and asked to develop the present report. Upon reviewing the terms of reference and defining a feasible scope of work, the panel noted that

recent documents (some of them produced after the terms of reference were created) already provide a reasonable architecture-level survey of the Defense Department's M&S, as well as a vision statement. In particular, the Office of the Secretary of Defense's (OSD's) Defense Modeling and Simulation Office (DMSO) has developed a substantial Master Plan for M&S, the purpose of which is to establish a common technical framework for DOD's M&S.[1] Given this body of existing material, the panel focused its efforts on key issues that have previously received little or insufficient attention. The objectives the panel set for itself, then, were (1) to clarify why the Department of the Navy leadership should care and be concerned about the *substantive content and comprehensibility* of M&S; (2) to assess what the Navy Department (and DOD) may need to do to benefit fully from the opportunities presented by M&S technology; (3) to clarify what M&S can and cannot be expected to accomplish in aiding decisions on technical, force-composition, and operations planning issues; and (4) to present priorities for M&S-related research.

The panel made no attempt to conduct a full survey of M&S relevant to the Department of the Navy. Much of the report deals with large-scale joint models such as those used in campaign planning, the evaluation of systems and new doctrinal concepts, or joint training—e.g., M&S such as the Joint Warfare System (JWARS) and the Joint Simulation System (JSIMS) systems now under development. The report has less to say about engineering- or engagement-level models, although it discusses the important role of simulation-based acquisition. Finally, this report is not a "forecast," nor does it lay out "roadmaps" for what should be done decade by decade for the next 40 years. Instead, the panel has chosen to focus on a chronic problem that took many years to develop and will take many years to deal with effectively—the lack of a good military-science research foundation on which to base the modeling and simulation that it so much depends on—and on priorities for remedying that problem over the years ahead.

Panel membership included experts in the research for and development and application of modeling and simulation, in both defense and nondefense domains. It also included experts in force planning; operations planning; applied mathematics, including probability and statistics; modeling and simulation theory; physics, including statistical mechanics; control theory; computer science; electrical engineering; operations research; gaming; and strategic planning.

The panel met eight times to receive briefings from Service and industry

[1]See Defense Modeling and Simulation Office. 1995a. *Department of Defense Modeling and Simulation (M&S) Master Plan*, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C., October; Kaminski, Paul G., Under Secretary of Defense for Acquisition and Technology. 1996. "DMSO 'Modeling and Simulation,'" Keynote address at DOD Fifth Annual Industry Briefing, Alexandria, Va., May 22; and other materials—both formal and informal—available from the DMSO or the DMSO's World Wide Web site at http://www.dmso.mil.

representatives, visit facilities, deliberate, and draft its report. It also participated in the three plenary meetings for the overall study. The first plenary meeting, in March 1996, established organization and a common starting point for the entire study. It included presentations by the Chief of Naval Operations and other high-level officials of the Navy Department, the other Services, the Defense Department, and industry. The subsequent plenaries were for drafting, comparison and integration across panels, the working out of cross-cutting issues, and synthesis (reflected primarily in *Volume 1: Overview*). The result follows. The report (which consists of a summary, the main report, and a set of appendixes) discusses modeling and simulation as a foundation technology for many developments that will be central to the Department of the Navy and Department of Defense over the next 3 to 4 decades.

The panel report is, of course, a product of the whole. However, the Vice Chair, Paul Davis, organized and led report preparation. He and Richard Ivanetich also compiled the panel's work and briefed it to study leadership along the way.

Acknowledgments

The Panel on Modeling and Simulation is indebted to many people who provided briefings, scientific papers, or discussion time. It gives special thanks to Darryl Morgeson and Chris Barrett of Los Alamos National Laboratory, Jeff Grossman of the Naval Research and Development Laboratory (NRaD), Les Parrish of SPAWAR, Bill Stevens and Jeff Steinman of Metron, Inc., Henson Graves of Lockheed Martin, Tom Skillman of Boeing, Timothy Horrigan of Horrigan Analytics, CDR Dennis McBride, USN, of the Office of Naval Research, and Wayne Hughes of the Naval Postgraduate School. CAPT Jay Kistler, USN, was the panel's contact with the Navy, the study's sponsor. Both he and CDR McBride provided useful briefings and contacts.

The panel also acknowledges RAND's courtesy in supplying several of the figures used to illustrate concepts discussed in the report.

Contents

Executive Summary

OVERVIEW: CONCLUSIONS AND RECOMMENDATIONS

The Department of the Navy must take a new look at modeling and simulation (M&S):

- The very nature of warfare is changing, perhaps drastically. The U.S. style of war is becoming technologically complex and dependent on distributed and interconnected systems. Modeling and simulation will be core tools for planning and conducting warfare as revolutionary changes in military affairs take place, especially since intuition based on past wars will become less helpful over time.
- Indeed, *independent of Navy and Marine actions*, M&S will be deeply embedded within joint command-and-control systems. Without enhanced efforts, the Navy and Marine Corps will not understand the strengths or limits of such models and simulations, nor be proficient with them.
- M&S will also become a core feature of system development and acquisition, as is the case already in leading-edge civilian industry. Because of its centrality, M&S should be seen as an enterprise technology in itself—part of the revolution in business affairs that is now a key element of the Department of Defense's (DOD's) overall strategy.

While the future of M&S should be exceedingly bright, the Department of the Navy will not be able to exploit its potential unless it attends to serious and chronic shortfalls—the most important of which relate not to software, but to the quality and content of the underlying models. Dramatic advances are being made

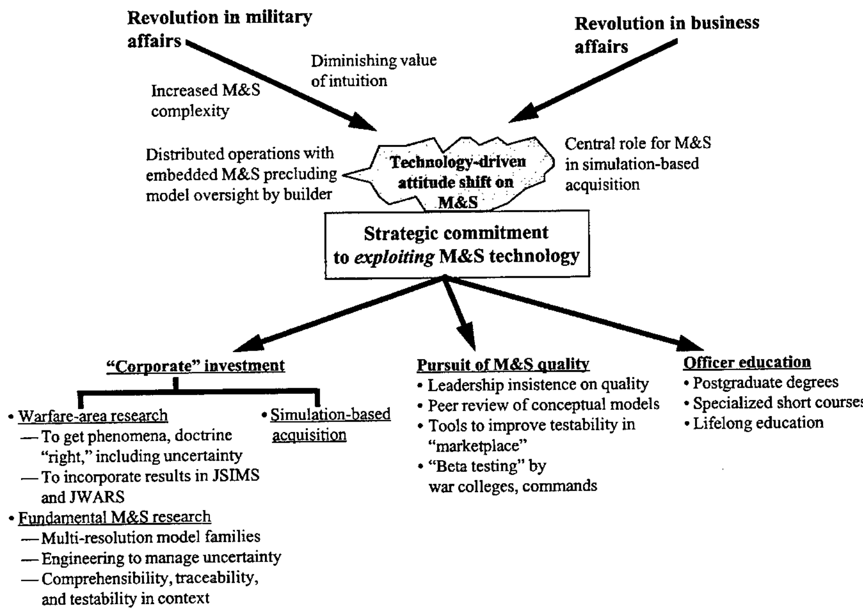

FIGURE ES.1 Looking at M&S strategically.

in DOD's M&S, but these advances are associated mainly with computer and software technologies. By contrast, too little attention has been focused on the content of the models themselves, or on the research base needed to create that content. Failure to address this shortfall will inevitably lead to less effective but more expensive combat forces and—quite possibly—serious operational failures. The escalating complexity of planned systems and operations creates profound integration challenges requiring superb M&S for success—and for the avoidance of downright failures.

All of this suggests that the Department of the Navy needs to make an attitude shift regarding M&S, which has never previously merited a high priority for leadership attention. Today, what is needed is a strategic commitment to exploiting M&S (Figure ES.1). This, of course, would lead to a strategy, policy, and investment actions. In this report the Panel on Modeling and Simulation identifies priorities for such matters. One priority involves the two principal joint simulation programs for training and analysis, the Joint Simulation System (JSIMS) and the Joint Warfare System (JWARS), respectively. *Since first-generation versions will be quite imperfect and the systems may last 10 to 20 years, the panel recommends that the Department of the Navy take an active role to ensure that JSIMS and JWARS are produced as evolving systems that incorporate future research results.* This is not simply a management issue, but rather something very challenging technically, since the architecture of the simulations

must allow this evolution. Also, links must be created between the research and M&S worlds, links that do not now exist. And the research itself must be expanded substantially.

The research needed falls into two categories, warfare-area research and fundamental research. The former involves understanding the processes that take place in combat (e.g., those of littoral operations or long-range precision strike). This is military science, not programming, and commercial industry will not lead the way. The fundamental research needed involves theory and methodology that will make it possible to design and construct sound and reusable models and simulations that will also be comprehensible, flexible, and testable in specific contexts. Achieving these features will require major advances. *It follows that the panel recommends that the Department of the Navy (working with other Services and DOD as appropriate) establish a robust but focused program in research, with both warfare-area research and research on fundamental theory and methods.*

Unfortunately, research alone is not sufficient, and its fruits are often not harvested because researchers, M&S developers, warfighters, and other leadership figures are often disconnected. *Accordingly, the panel recommends that the Department of the Navy establish processes that ensure early scientific review of models emerging from research, a competitive atmosphere in which "the market" of model users is both encouraged and assisted in constantly evolving their M&S to represent the best available knowledge (i.e., in assimilating improvements), and a general emphasis on quality, including the ability to represent uncertainty. Accomplishing this will require a multiyear commitment of senior leadership, because the baseline culture is very different from the one needed.* Elements of a changed approach would include enhanced officer education and continual "beta testing" of models and simulations by organizations such as the war colleges and commands, testing that extends deeply into content, not merely software performance. However, such "operators" will need substantial assistance from the scientific community.

In addition to addressing the quality and content of models and simulation, the Department of the Navy needs to make investments that cut across usual organizational stovepipes and budget accounts. This relates to the promise of simulation-based acquisition. *Accordingly, the panel recommends that the Department of the Navy treat simulation-based acquisition (SBA) as a key enabling technology with extraordinary long-term leverage and that it organize and invest consistently with that enterprise-technology view.*

RICH OPPORTUNITIES FOR MODELING AND SIMULATION

Modeling and simulation (M&S) offers the promise of greatly enhancing future naval force capabilities and achieving major cost savings. To cite a few examples:

• *Concept development.* Next-century naval forces will require new operational concepts and force structures—in some cases radically different from current ones. Ours is an era of military ferment analogous to that of the 1920s or 1930s. M&S can help screen, design, and test concepts and force structures before irrevocable commitments to them are made. In principle, M&S could also be compelling enough to "force" needed changes of doctrine before disasters occur in war.

• *Simulation-based acquisition (SBA).* Representations of proposed system designs can be constructed and tested in simulated environments. These virtual prototypes can be used to refine system requirements and relate tradeoff and engineering decisions to these requirements. Subsequently, computer-based representations can be maintained as development and production occur, and as modifications are introduced throughout the life cycle. The results can be more affordable systems that are better attuned to an operator's needs, easier to assimilate, and easier to modify. The remarkable success of the Boeing 777's development merely tapped the surface of what will eventually be possible.

• *Decision support.* M&S can assist commanders in their planning for combat and other military operations. Key uses include developing and assessing proposed courses of action, mission planning and rehearsal, and dynamic situation assessment and adaptation in the course of battle. M&S-based decision support also has an important role in peacetime activities such as concept evaluation and resource allocation.

Models and simulations are being used for all of these functions today. Indeed, the breadth of M&S is enormous, as suggested by Figure ES.2. The best-known recent successes have been in training, but there is rapidly growing documentation on valuable applications throughout Figure ES.2's cube—as judged not only by those advocating M&S, but also by senior commanders. Documented examples of recent high-leverage payoffs from M&S are given in Chapter 2, but the future is more relevant in this study. It is especially significant for the future that M&S will be thoroughly embedded in command-and-control systems. There have already been numerous exercises in which differences between the real and the simulated have been blurred or made invisible to some participants. This will be increasingly the case as U.S. forces adopt the concepts sketched in the Joint Chiefs of Staff's *Joint Vision 2010* (Shalikashvili, 1996).

THE POTENTIAL FOR FAILURES AND DISASTERS

Cautions Amidst Enthusiasm

Despite this bullish introduction and the fact that M&S will surely "take off" in the commercial sector, the potential of M&S for the Department of the Navy and DOD may not be realized in the foreseeable future. Some of the DOD's most

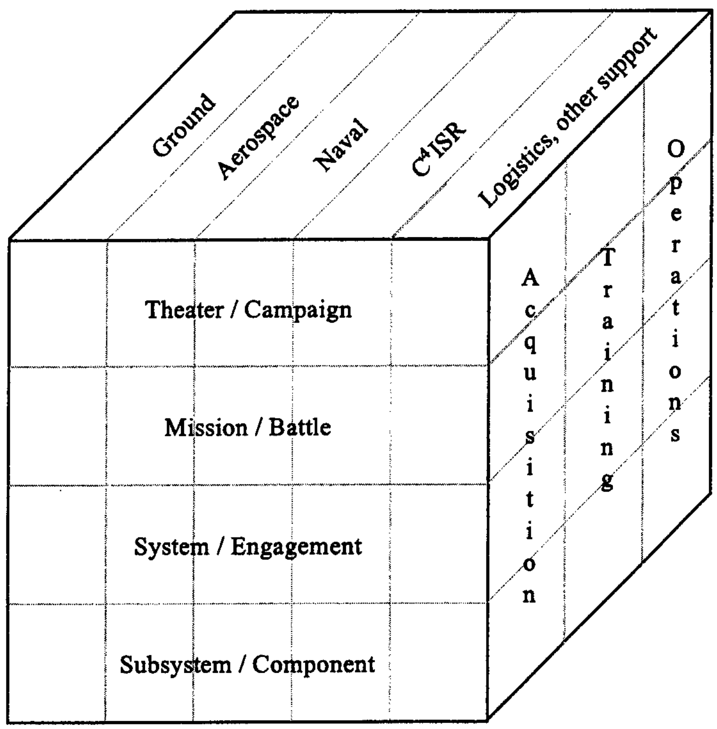

FIGURE ES.2 The scope of DOD's M&S activities. SOURCE: Adapted from Defense Modeling and Simulation Office (1996d), Figure 2-1.

important and expensive M&S efforts may fail or—perhaps worse—end up saddling the DOD components with mediocre, inflexible, and sometimes misleading tools that impede innovation and improvement of content. There is also the potential for disasters due to overdependence on M&S for images and predictions that appear more valid than they actually are (the "Spielberg" effect). This could cost lives and undercut military operations.

One basic problem is that DOD's investments have been concentrated more on content-neutral computer and software technologies rather than on the content of the models.[1] On the technology side, much has happened, for example, object-

[1]There are some partial counterexamples, of course: e.g., "knowledge acquisition" for the "semi-automated forces" (SAFOR) used in the Joint Countermine Operational Simulation (JCOS) and the Army's tactical training system, new modeling of C^4ISR effects in JWARS, development of object models, and research on configural effects. Even in these cases, however, the work has often been more like "computer modeling" than establishing an empirical and theoretical research base.

oriented programming, high-performance computing, computer-aided design, and establishment of the communications protocols and infrastructure for everything from collaborative, distributed, multidisciplinary, simulation-based engineering design to large-scale distributed war games. The dramatic progress in technology will assuredly continue because of commercial developments and DOD efforts such as those of its Defense Modeling and Simulation Office (DMSO).

An Undernourished Knowledge Base

In contrast, there has been curiously little investment in the knowledge base determining the substantive content and quality of much M&S—particularly higher-level M&S needed for mission- and campaign-level work. It is an open secret, and a point of distress to many in the community, that too much of the substantive content of such M&S has its origin in anecdote, the infamous "BOGSAT" (bunch of guys sitting around a table), or stereotypical versions of today's doctrinally correct behavior. There is a need for focused research on the phenomena of combat and other military activities, both historical and prospective. This is the realm of military science.

Another shortfall in knowledge relates to theories and methods for conceiving, designing, and building *models* (as distinct from software). Symptoms of the problem are evident if one observes that DOD's M&S often consists of nothing more than the computer code itself: there is no separable documented "model" to be reviewed and improved, nor any way to readily understand the assumptions generating the simulation's behavior. This can hardly be a comfortable basis for decision support.

Complex Systems

The inherent complexity of the systems and force operations that DOD is attempting to simulate introduces new difficulties. Too many forecasts are extrapolating unreasonably from the Boeing 777 experience, and from M&S successes in weapon-system and small-unit training, to imagined M&S systems of extraordinary complexity. Recent failures such as the automated Denver airport baggage system and the Federal Aviation Administration advanced air-traffic-control system suggest the difficulties associated with reliably modeling and engineering complex systems. The military operations envisioned for future forces in the information era involve exceedingly complex systems.

Complexity is a multifaceted concept, but to appreciate some of what is involved here, the panel notes that planners, commanders, and engineers are most familiar and comfortable with systems (and models) that are primarily static, linear, and deterministic. However, automated and integrated military systems (the systems-of-systems approach) involve systems and models that are dynamic, nonlinear, and heterogeneous interconnections of mixed subsystems, with a much

more sophisticated treatment of uncertainty, including uncertainties about the opponent's intentions and actions.

The panel's principal observation here is that dealing with such complex systems effectively is a decades-long task. In the meantime, there is need for humility, multiple approaches and competition, patience, and hedges. And even long-run success will require profound changes in the way M&S is conceived and designed, as well as an across-the-board attention to its content and validity—in a sense that suitably recognizes uncertainties. The panel describes what is needed below.

Assimilation of M&S Technology

Finally, there are the problems of assimilation and exploitation. M&S is an enabling technology, but its value is cross-cutting, and it has no natural single home. Nor should it, since—in a partial defense of stovepiping—the majority of work must be dictated by the needs of individual applications. However, whenever such a cross-cutting technology is introduced or its use expanded, there are organizational and managerial challenges. The key to success is often having a strategy embraced by the organization's leadership.

Against this background diagnosis, the panel has observations and suggestions in each of a number of subjects. In essence, the recommendation is for a concerted and long-overdue effort to improve the research base for the Department of the Navy's and DOD's models, and to ensure that the results of research are in fact incorporated. The panel's recommendations include priorities and suggestions for a strategy, not simply general funding of research.

RECOMMENDATIONS ON JOINT MODELS

Concerns

One useful focus for the Department of the Navy's thinking about M&S is the set of joint systems now in development (most prominently JSIMS and JWARS). Taken together, these worthy programs (including the Service components) have a price tag approaching $1 billion. It is DOD's intention that JSIMS and JWARS will become the core for all future joint work on training and analysis, respectively. Although this is unlikely (a wider range of models will probably prove necessary), it may indeed be that JSIMS and JWARS will dominate the joint M&S scene for the next 20 years. Thus, it is important to the Department of the Navy that naval forces be adequately represented. Otherwise, valuable training opportunities will be compromised and the Navy and Marines will suffer in the competitions about doctrinal changes, future missions, and force-structure tradeoffs. More generally, the quality of joint work will suffer.

Unfortunately, it is likely that first-generation versions of JSIMS and JWARS

will not be satisfactory—even with heroic efforts and even though the products will have many excellent features. There will be major shortcomings with respect to both content and performance. *Consequently, the panel recommends that the Navy insist that DOD and the program offices adopt open-architecture attitudes that will promote rather than discourage substitution of improved modules as ideas arise from the research and operations communities, and that they build explicit and well-exercised mechanisms to ensure that such substitutions occur.*

This recommendation may seem uncontroversial, and it calls for no more than what some of the programs (notably JSIMS) are projecting on viewgraphs, but the history of DOD modeling has often been to produce relatively monolithic and inflexible programs. Further, there has been great DOD emphasis in recent years on avoiding alleged redundancies, collecting "authoritative representations," and exercising configuration control. The panel observes widespread frustration among analysts and other substantive users of models, who see DOD's M&S efforts as driven by civilian and military managers who think models are commodities to be standardized, who sometimes seem to value standardization more highly than quality (harsh words, but too important to be omitted), and who have given near-exclusive emphasis to software technology issues. They and the panel believe that M&S should instead be seen as organic, evolving, and flexible systems with no permanent shape (but with standardized infrastructure, including many component pieces).

In fact, the visionary technical infrastructure being promoted by OSD's Defense Modeling and Simulation Office (DMSO) (and software technologists) will permit the open system approach and will permit competition among alternative models (e.g., alternative representations of ballistic-missile defense, mine warfare, or command, control, communications, computing, intelligence, surveillance, and reconnaissance (C^4ISR)). Thus, while it would be easy for JWARS, JSIMS, and other systems to end up as rigid monoliths, with the right architecture and organizational structure DOD can have its cake and eat it: it can have "standard configurations" while still making it easy for users to substitute model components as new ideas and methods emerge. An important but more subtle aspect of this infrastructure is connecting model evolution to the R&D and operational communities concerned with both current and futuristic doctrine; and, significantly, nurturing a competition of ideas and models. In that way the evolution will be more like survival of the soundest than like continuation of what has previously been approved.

The panel underlines the problem of incorporating research results when they exist because, at present, the communities that do research and the programming of models often do not communicate well and there is little pressure to assure that the "best" models are reflected in M&S. Indeed, there is much pressure to avoid changes.

Technical Attributes Needed in Joint Models

Against this background of concerns, *the panel recommends that the Navy advocate an approach to joint-model development that has a long-haul view and an associated emphasis on flexibility.* The groundwork should be in current model-building efforts for the following, which will be important in selected applications in the years ahead:

- *Multi-resolution modeling*, not only of entities, but also of physical and command-and-control processes, with the objective of building integrated models of families with different levels of resolution.
- *Explicit decision models* representing the reasoning and behavior of commanders and different levels, and reflecting in natural ways courses of action, plans, and the adaptations that commanders make in the course of operations.
- *Diverse representations of uncertainty*, including use of probability distributions (and, sometimes, alternatives such as fuzzy-set concepts), even in aggregate-level models.
- *Systematic treatment of important correlations* (e.g., the "configural effects" of mine warfare and air defense).
- *Explanation capabilities* linking simulated behavior to situations, parameter values, rules and algorithms, and underlying conceptual models.
- *Mixed modes of play* that are interactive, selectively interruptible (e.g., for only higher-level commander decisions), and automated. (The panel regards the option for human play as critical for analytic applications as well as training, and the option of closed play, e.g., of the opponent, as critical for training.)
- *Testing of new doctrinal concepts* requiring new entities, attributes, and processes.
- *Different types of models.* The systems should accommodate model types as diverse as general state-space and simple Lanchester equations, entity-level "physics-based" models, and agent-based models with emergent behaviors. They should employ such varied tools for such uses as statistical analysis, generation of response surfaces, symbolic manipulators, inference engines, and search methods (e.g., genetic algorithms). The models must be able to deal not just with old-style head-on-head attrition warfare, but also with maneuver warfare on a nonlinear battlefield in the information era, and with operations in urban sprawl. They must reflect different command-and-control concepts.
- *Tailored assembly.* The systems should facilitate tailored creation of models, including relatively simple M&S for specific applications. That is, one should conceive of JSIMS and JWARS as tool kits with rapid-assembly and modification mechanisms. Excessive complexity obfuscates and paralyzes.

In some respects, the last item is the most important. Given the breakthroughs in software technology over the last two decades, it is feasible (though

not easy)—and essential—for major M&S efforts to be designed for frequent adaptation, specialization, and module-by-module improvement. One should think of assembling the right model, not taking it from the shelf whole. Further, it should be possible to discard or abstract complexities irrelevant to the problem at hand. Doing so runs directly counter to the common inclination to seek high resolution for everything, but tailored simplifications are crucial in applications—especially when they involve "exploratory analysis" over diverse situations and assumptions rather than point calculations. The type of analysis is crucial when uncertainties are large—as they often are. In any case, the need for simplicity is generally much better understood by those who have conducted studies or exercises, or designed decision-support systems, than by those who develop software.

The panel notes, however, that there are limits to what can be accomplished by assembly or composition. The Department of the Navy and DOD should be skeptical about the notion that a single system (e.g., JWARS or JSIMS) will prove useful to a wide range of communities. It is one thing to assemble and tailor components for one study rather than another, or for one exercise rather than another. It is a very different matter to have the same system and library of components support a broad range of different functions (testing, exercises, force planning, etc.). Viewgraphs postulating such versatility do not constitute an existence proof.

RECOMMENDATIONS FOR RESEARCH

Research in Key Warfare Areas

As noted above, there has been relatively little recent investment in understanding the phenomenology of military operations at the mission and operational levels. Much of the basis for related M&S is still programmer hypothesis and qualitative opinions expressed by subject matter experts. This has not always been so. During and after World War II, operations research worked from a rich empirical base, but now the United States is entering a period of nonlinear, parallel, information-era warfare for which the intuition of scientists, operations researchers, and warriors is insufficient and unreliable. Further, it will be relying on complex systems working as designed in multifaceted joint campaigns. Success may be much less tolerant of errors in concept and execution than in days past. Indeed, some of the doctrinal concepts under discussion will involve very high risks.

Given, then, that improvement of the research base is essential, how might it be accomplished? Rather than merely urging general support for research, the panel recommends a managerially focused approach with priorities and mechanisms for assuring relevance and assimilation. *Accordingly, the panel recommends that the Navy and Marine Corps select a few high-priority warfare areas*

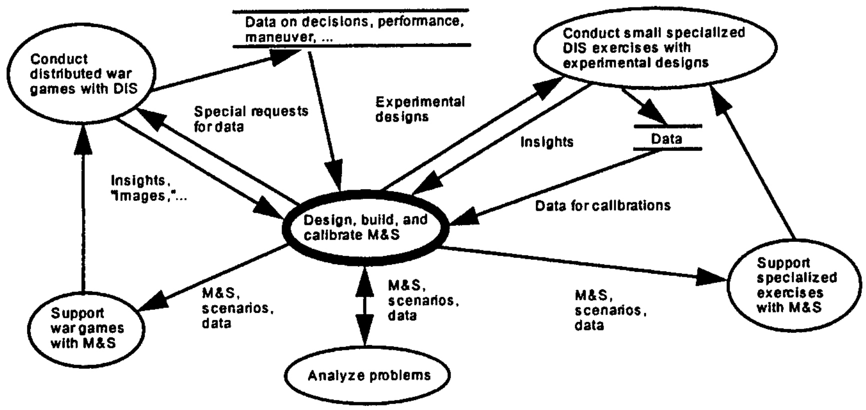

FIGURE ES.3 Using exercises as a source of empirical data for M&S. SOURCE: Reprinted, by permission, from Davis (1995b). Copyright 1995 by IEEE.

and create research programs to support them. These programs should be organized so as to ensure close ties to operational and doctrinal-development communities, and to relevant training and exercise efforts that could be mined as a source of empirical knowledge (e.g., as suggested in Figure ES.3, which would exploit emerging capabilities for distributed interactive simulation). This is a nontrivial and potentially controversial suggestion, since the long-standing tradition has been to avoid—and even prohibit—extensive data collection for use beyond those being trained. The costs of such efforts would be small in comparison with those for buying and operating forces, or even procuring large models. Although the Department of the Navy (and DOD) need to make up for past failures to invest adequately in research, this is a domain in which a total of $20 million to $30 million per year can accomplish a great deal.

As a first list of warfare areas for focused research, the panel recommends the following, which have some overlaps:

- Expeditionary warfare and littoral operations;
- Joint task force operations with dispersed forces;
- Long-range precision strike against forces employing countermeasures;
- Theater-missile defense, including counterforce and speed-of-light weapon options, against very large ballistic-missile and cruise-missile threats; and
- Short-notice, early-entry operations with opposition.

Each of the above warfare areas has major knowledge gaps that could be narrowed by empirical and theoretical research closely tied to the "warrior communities."

This report describes key attributes of research programs for such warfare areas. An overarching theme is the need to take a holistic approach rather than

one based exclusively on either top-down or bottom-up ideas. A second theme is that the research should be seen as focused military science, not model building per se. This will determine the type and range of people involved, and also the depth of the work.

Two examples may be useful here. The first is the challenge of developing command-control concepts for highly dispersed Marine Corps forces operating in small units far from their ship-based support and dependent on a constellation of joint systems. The Marine Corps is studying alternative concepts in the Hunter/Warrior experiments. Such experiments need to be accompanied by systematic research and modeling of different types, perhaps including new types of modeling useful in breaking old mind-sets. It is plausible, for example, that cellular-automata models could help illuminate behaviors of dispersed forces with varying command-control concepts ranging from centralized top-down control to decentralized control based on mission orders. To its great credit, the Marine Corps is exploring such possibilities, opting to accept some "hype and smoke" in the realm of controversial complex-system research in exchange for new perspectives and tools useful in doctrinal innovation. While the panel does not believe such simplified models will prove adequate in the long run, they can be very helpful in developing new hypotheses.

A Navy example involves mine and countermine warfare. From prior research based on sophisticated probabilistic modeling accounting for numerous "configural effects" (i.e., effects of temporal and spatial correlations), it is known that effective strategies for mine-laying or penetrating minefields are often counterintuitive. By exercising such models and simulation-based alternatives in an exploratory manner (as distinct from answering specific questions), it should be possible to develop decision aids of great value in training, acquisition, and operations. Such aids should not, however, focus only on "best estimate" single-number predictions; they should instead provide commanders with information about odds of success, as a function of information. If the aids are to be useful, they must be informed by an intimate understanding of operational commanders' needs.

Recommendations on Fundamental Research

While many research activities are best driven by applications, other critical areas of M&S research require more fundamental research that might be sponsored by the Office of Naval Research (ONR), other Service analogues, the Defense Advanced Research Projects Agency (DARPA), and the Director of Defense Research and Engineering (DDR&E). In what follows the panel suggests particular subjects for fundamental research, divided into theory, advanced methodologies, and infrastructure (including tools).

Modeling Theory

The panel recommends that research on the following be given priority:

- *Multi-resolution modeling, integrated families of models, and aggregation-disaggregation.* Multiple-resolution modeling depicts phenomena at different levels of detail. In some cases a single model can operate at different levels of resolution with appropriate kinds of consistency. More often it is possible—although unusual and difficult—to design integrated families of models that can be mutually calibrated using information available in many forms and resolutions, and to do so with full recognition of statistical averaging issues. Such families of integrated hierarchical models would be invaluable in all application areas and would substantially improve validity, traceability, and the design of exploratory analyses. How to design and build such multi-resolution models or families, however, is a frontier problem in modeling theory. A related subject is often called aggregation-disaggregation. This often refers to distributed simulations in which, within the course of a simulation, some of the entities must be disaggregated and reaggregated (e.g., a Marine Corps company might have to be disaggregated to engage simulated entity-level opponents and then reaggregated to continue its maneuver). While this is relatively straightforward from a software perspective, there are deep questions concerning the consistency of behaviors at different resolution levels.
- *Agent-based modeling and generative analysis.* Some of the most interesting new forms of modeling involve so-called "agent-based systems" in which low-level entities with relatively simple attributes and behaviors can collectively produce (or "generate") complex and realistic "emergent" system behaviors. This is potentially a powerful approach to understanding complex adaptive systems generally—in fields as diverse as ecology, economics, and military command-control. A fundamental step in developing particular models and simulations is deciding which attributes and interactions to represent, and in what detail. This choice should be the one that most adequately describes the phenomena one is trying to observe, but that choice is often not known until the subsystems are connected and the simulation is run. Thus, methods should be developed to allow one to iterate on the choice of the initial representations of subsystem models, based on results of their use in interconnected systems.
- *Semantic consistency.* Phenomenological representations in different simulations need to interact with one another in distributed simulations. Such interaction is meaningful only if the representations are "semantically consistent," that is, if there is a shared understanding of what concepts and data "mean." This requires commonality of context and definition (or well-understood translations). The Navy should track related research, adding to it for special purposes. It should also support DMSO efforts to develop common models of the mission

space (CMMS), which will assist in establishing semantic consistency in particular contexts and in developing integrated families of models.

Advanced Methodologies

The general task of developing and using models and simulations, and the particular activity of forming phenomenological representations, would be aided by methodological advances. Particular topics are as follows:

- *Characterization of uncertainty.* No matter how careful one is in preparing for a simulation, certain attributes and interactions will have some measure of uncertainty. Often, uncertainties dominate the problem. Methods to track the propagation of uncertainties should be developed since they can lead to large uncertainties in the output of the simulation. This is a particular challenge in heterogeneous, nonlinear dynamical systems, where uncertainties in components can interact in nonintuitive and unpredictable ways. The so-called "butterfly effect" in chaotic systems is a well-known popular example.
- *Exploratory analysis under uncertainty.* Running a simulation for one set of fixed conditions is generally not satisfactory since there are often large uncertainties throughout the system. Even normal sensitivity analysis on a one-variable-at-a-time basis does not suffice because of interaction effects. An important research area, then, is developing ways to use modern computer power to explore the space of simulation outcomes and to search for interesting regimes (e.g., regimes representing high or low risks for an operation or for especially profitable, or unacceptable, performance of a weapon). This research has implications for the design of models (some of it closely related to multi-resolution modeling), search engines, and visualization methods. It has even more profound implications for analysis and decision making because it encourages decision makers to ask not about best-estimate outcomes, which are often no more likely than very different ones, but rather about how outcomes of a strategy would be likely to vary as a function of the many assumptions in "scenario space." This can help by focusing attention on the need to avoid "dangerous regimes" in the course of operations, by focusing attention on the search for crucial information, and by emphasizing the need for both hedging and adaptability. This approach, of course, is quite different from the search for mythical optimality.

Infrastructure, Tools, and Supporting Technology

With regard to infrastructure, the following research areas are of particular significance:

- *Intellectual infrastructure.* Scientific and engineering disciplines typically have a mathematical language in which to frame and solve their problems—

e.g., the use of calculus for disciplines as diverse as aeronautical engineering and chemistry. In contrast, there is no widely understood and adopted theoretical basis for M&S. To some extent, object-oriented modeling (not programming) is helping here, but in practice it usually deals with only some of the problems. While mathematics and systems theory can form a common language, modeling assumptions and their consequences tend to be domain-specific and implicit. Even worse, the only underpinning to many simulations is the computer code in which they are written. To help create the needed intellectual infrastructure, the Department of the Navy and DOD should cooperate with industry and universities in encouraging the development of theory and the promulgation of standard texts and case studies. DOD's adoption of software engineering methodologies (e.g., in the JWARS effort) is useful here. It may also be useful for the Department of the Navy, other Service components, and OSD to cooperate in developing "virtual centers" exploiting the World Wide Web, and in establishing additional peer-reviewed journals and scientific conferences overseen by research institutions.

- *Object repositories and interface standards to enhance reusability and composability.* Object-oriented technology admits the possibility of assembling major parts of simulations to meet the demands of a particular application from sets of stored objects representing entities and processes. Realization of this capability requires being able to manage large numbers of objects and to ensure consistency despite involvement of multiple developers. Such a capability could reduce costs in simulation development and allow flexibility in simulation application. A key of OSD's strategy in this domain is embodied in the high-level architecture (HLA), which establishes standards for M&S that may be used in federations employing distributed simulation. Despite controversy and anxiety about program costs, the HLA is a needed step in the direction of increased modularity and interoperability. It will have many long-term benefits. The Navy should support and exploit the HLA initiative—recommending modifications as needed.
- *Explanation/traceability capability.* This capability applies to all phases of the management process. For example, it would help document the source code with multimedia techniques so that one could understand the phenomena being represented, and it would help explain the results of a simulation by displaying the logic trail that led to the results. Realization of this capability would figure centrally in achieving the verification, validation, and accreditation (VV&A) of simulations, both in the formal sense and to the satisfaction of individual users. This capability is important for field commanders, managers, and engineers.

For example, commanders using M&S to assess courses of action may need to know the following: On what assumptions do the simulation outcomes depend critically; should those assumptions be modified and the simulations rerun; what if the component commanders are given contingent orders? The implications of having both comprehensible models and effective explanation/traceability capa-

bility are so significant for VV&A that the Navy should undertake more general efforts. The Navy should use commercial products where appropriate and should foster commercial development where necessary, since the capabilities required will have more general value (e.g., to computer-aided design). Such commercial coupling is necessary because development of these capabilities will be expensive.

- *Other tools.* Many other tools are badly needed. These include tools for (1) automated scenario generation and experimental design and (2) postprocessing and data analysis.

Research Is Not Enough: Planning to Incorporate Its Fruits

There is one further challenge associated with research: assuring that its products are recognized and used appropriately. This is challenging because the research, M&S-building, and user communities are reasonably distinct. Those building M&S often are only minimally acquainted with cutting-edge research in either the phenomenology of warfare operations or modeling methodology. Further, their sponsoring organizations are often more interested in the stability of M&S (and the related gargantuan databases) than in improvements of "theory," the rewards of which may be less than immediately tangible. This state of affairs probably continues because so little higher-level M&S is bounced against empirical realities. The panel has several suggestions here:

- Requiring documentation of "conceptual models" (as distinct from details of the implementing programs and databases),
- Providing scientific review of such models to advise the Navy about the quality of the models in relation to scientific knowledge and best practices in the community, and
- Redefining the JWARS and JSIMS programs to have a continuing component responsible for reviewing, sponsoring, and incorporating research results. The function should be one of nourishing military science, not merely administration or auditing. Major changes are needed if there is to be a resurgence of in-depth study of military phenomena and the kind of open scientific discussion and debate that will lead to top-quality M&S.

RECOMMENDATIONS ON ASSIMILATING AND EXPLOITING M&S

The Need for Strategic Commitment

Finally, there are the organizational problems of assimilation and exploitation. As noted above, M&S is an enabling technology. However, whenever such a cross-cutting technology is introduced, there are organizational and managerial challenges. It is commonplace for the organization to measure the value of

investments against the wrong yardsticks (e.g., saving money in narrow domains, as distinct from changing the very way the organization does business and improving effectiveness for mainstream missions). It is also common for investments to go awry because (1) the new technology is procured and used as an add-on without sufficient buy-in and influence by the organization's core responsibility and workers, (2) too much is done by committee without leaders and champions who understand the core business, or (3) the educational groundwork has not been laid. Despite much ongoing success, all of these M&S-related problems are visible in DOD's components, including the Navy and the Marine Corps.

Based on the history of technology assimilation and the specifics of the current situation with respect to M&S, the panel recommends that the Department of the Navy make a strategic commitment to the success of exploiting M&S. Such a commitment would have consequences for organizational structure and responsibility (although the panel makes no recommendations on such matters), investment mechanisms (e.g., assuring that investment funds are available without forcing program managers always to make tradeoffs within their own domains), and the establishment of clear policies and strategies that would make manifest the leadership's *demand* for constant improvements in "validity" (as understood in the context of sometimes-extreme uncertainties) and usefulness to decision makers. As discussed above, the panel believes that the appropriate strategy would place considerable emphasis on warfare areas and cross-cutting modeling challenges, rather than still more emphasis on computer and software technology. To put this more bluntly, if funding tradeoffs are needed within M&S budgets, then the panel recommends giving higher priority to research improving model content rather than programming or reprogramming of current models.

In addition to investing in research to improve the quality and content of models, the Department of the Navy must organize, plan, and invest strategically if it is to enjoy the potentially great benefits of simulation-based acquisition. In doing so, it should take a long view because, as in other aspects of M&S, there are substantial obstacles. Success will be evolutionary over a period of decades.

Education for Next-generation Officers

One element of a strategy should be increased education in M&S for next-generation officers. The effective exploitation of M&S depends on the experience, knowledge, and wisdom of its practitioners, hence upon their education. *The panel recommends increased Navy investment in such education at all levels: for those who acquire and design M&S tools, and also for those who rely on them to guide acquisition, training, and operations.* Some of the education should be in the form of enhanced master's and Ph.D.-level programs. Other aspects should include short courses tailored for officers needing refresher courses, technology updates, and preparation for next assignments involving M&S management.

1

Introduction

BACKGROUND AND OBJECTIVES

This report discusses modeling and simulation (M&S) as a foundation technology for many developments that will be central to the Department of the Navy and Department of Defense (DOD) over the next 3 to 4 decades. As discussed in the preface, the report is part of a larger National Research Council study, *Technology for the United States Navy and Marine Corps, 2000-2035: Becoming a 21st-Century Force*. The study's Panel on Modeling and Simulation was asked to assume a cross-cutting role and, specifically, to address issues of model quality and validity—a matter of growing concern as the Department of the Navy and DOD become increasingly dependent on M&S for activities as diverse as training, acquisition, and operations planning.

Terms of Reference

The terms of reference (TOR) for the overall study made the following request:[1]

> [T]he review should place emphasis on surveying present and emerging technical opportunities to advance Navy and Marine Corps capabilities. . . . [T]he review should include key military and civilian technologies that can affect Navy and Marine Corps future operations. This technical assessment should evaluate which science and technology research must be maintained in naval

[1]Letter from Admiral J.M. Boorda, USN, Chief of Naval Operations, to Dr. Bruce Alberts, President of the National Academy of Sciences, November 28, 1995. See Appendix A.

research laboratories as core requirements versus what research commercial industry can be relied upon to develop.

More specifically with respect to M&S, the TOR stated:

The study should review the overall architecture of models and simulation in the DoD (DoN, JCS, and OSD), the ability of models to represent real world situations, and their merits as tools upon which to make technical and force composition decisions.

Defining the Scope of Work

The panel interpreted its charge in light of other developments and judgments about what it could most usefully accomplish consistent with the spirit of the request. The panel concluded that recent documents provide a reasonable architecture-level survey of DOD's M&S, as well as a vision statement. In particular, OSD's Defense Modeling and Simulation Office (DMSO) has developed a substantial Master Plan for M&S, the purpose of which is to establish a common technical framework for DOD's M&S.[2] This involves establishing a common high-level simulation architecture (HLA), conceptual models of the mission space (CMMS), and data standards—items that the panel will discuss later in more detail.

Figure 1.1, adapted liberally from the Master Plan (DMSO, 1996d), indicates the breadth of DOD's M&S. Figure 1.1 highlights several facts. First, M&S is accomplished at many levels ranging from engineering subsystems up to full-scale wars. This report deals largely with higher level issues shaded in Figure 1.1. Second, M&S is a key element of work in distinct functional areas—notably training, acquisition, and operations planning. Third, there is M&S for each of the components of military capability, that is, ground forces, naval forces, and aerospace forces.[3] And, as indicated at the left side, there are other dimensions that might have been highlighted: the size and resolution of the M&S, the nature and degree of human participation, and so on.

Given this existing material, the panel chose to focus more narrowly on key issues that have previously gotten insufficient attention. The objectives, then, were as follows:

1. Clarifying why senior levels of the Department of the Navy should care and be concerned about the substantive content and comprehensibility of M&S.

[2]See DMSO (1995a), Kaminski (1996), and other materials—both formal and informal—available from the DMSO or the DMSO's World Wide Web site at http://www.dmso.mil.

[3]The Master Plan's figure is somewhat different. It shows the three functional areas as training, analysis, and acquisition. It also focuses on the sponsoring component of models rather than the domain they cover.

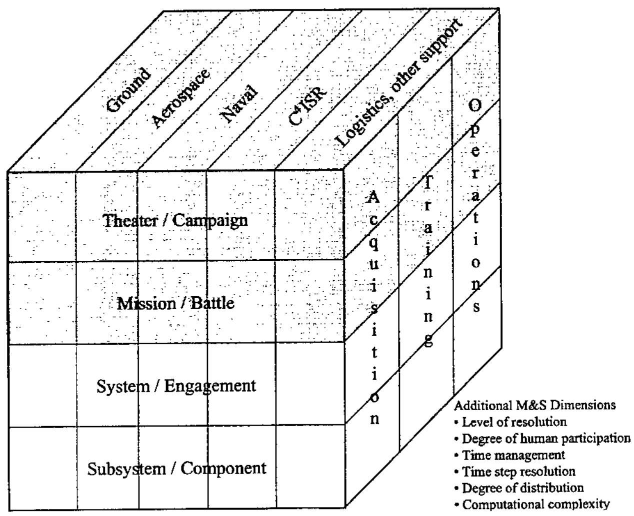

FIGURE 1.1 The scope of DOD's M&S activities. SOURCE: Adapted from Defense Modeling and Simulation Office (1996d), Figure 2-1.

2. Assessing what the Department of the Navy (and DOD) may need to do in order to benefit fully from the opportunities presented by M&S technology.
3. Clarifying what M&S can and cannot be expected to accomplish in aiding decisions on technical, force-composition, and operations planning issues.
4. Establishing what priorities should be for M&S-related research.

The panel also made no attempt to conduct a full survey of Department of the Navy-relevant M&S, given the enormity of the subject. Much of the report deals with large-scale joint models such as those used in campaign planning, the evaluation of systems and new doctrinal concepts, or joint training, for example, M&S such as the JWARS and JSIMS systems now under development. It has less to say about engineering- or engagement-level models. Examples also tend to focus on naval forces.

REASONS FOR THE DEPARTMENT OF THE NAVY TO BE INTERESTED IN AND CONCERNED ABOUT M&S

Top-level Reasons

With this background of definitions and distinctions, let us next ask why the leadership of the Navy and Marine Corps should be more than routinely interested in M&S. While M&S is already used throughout the Navy Department, this is not a good enough reason to justify special high-level attention. The reasons for interest, however, are several. They relate to

• Success of next-century naval-force visions (see, e.g., Johnson and Libicki, 1996);

• The size of the investment in M&S and the potential for saving money and improving effectiveness;

• Relationship between M&S prowess and Service competitiveness; and

• Obstacles to success in M&S, including problems of model quality and validity.

M&S as a Critical Factor

The panel's first observation is that the visions painted by the other panels of the larger study, *Technology for the United States Navy and Marine Corps, 2000-2035: Becoming a 21st-Century Force,* depend on extraordinary advances in M&S in the several domains. As one example, simulation-based acquisition (SBA) can bring about revolutionary changes in ship- and weapon-system development.[4] At the other end of the story, M&S will be an integral part of the command-and-control system that commanders use to develop and test operations plans, and to conduct mission rehearsals before going into battle. It will be an integral part of adaptive planning during campaigns—especially in the longer term when military adversaries become considerably more capable than they are today. If phrases like *dominant battlefield knowledge* (as distinct from a more static situational awareness) mean anything, they mean that future commanders will be able to project and predict (with an understanding of probabilities) and therefore adapt quickly and decisively.[5]

Economic Issues

A second reason for interest is economic. The Department of the Navy (and DOD) is investing large sums to develop aspects of M&S. The proper infrastructure creating interoperability and reusability of model components and data should lead to large dollar savings. However, there is no guarantee that this will occur or that naval forces will be adequately represented. Thus, the Department of the Navy has an interest in active participation and leadership with respect to M&S activities.

M&S and the Department of the Navy Competitiveness

A third reason is another practical one: the Navy's competitiveness in relation to the other Services will depend on the expertise that it develops in M&S.

---

[4]This was emphasized in a report of the Naval Research Advisory Committee (1994).

[5]For discussion of troublesome future adversaries, see Defense Science Board (1995). For discussion of information dominance, see, e.g., Johnson and Libicki (1996).

The latter will be so important to future joint operations that the Services having the best expertise and systems may have the competitive edge for roles and missions (e.g., command and control for a joint task force or theater missile defense), leadership positions generally, and budget share. The converse is also true: if the Department of the Navy is not sufficiently expert—and, even if it is, if it is not sufficiently "connected" to the joint-simulation world—then it should expect to suffer in the competition for budget shares.

Potential Roadblocks

This said, there are many potential roadblocks to success of M&S in the Department of the Navy (and DOD). The advances will not occur naturally, except in domains where the commercial sector is driving progress. For example, one can hardly expect the commercial sector to develop decision-support systems for commanders to use in war, with all the associated complexities, dynamics, competitive and lethal processes, and fog. Nor can one expect productivity enhancing and cost-cutting successes in the commercial world to automatically be assimilated into government organizations: generational changes of technology and reengineering are often painful. They do not happen without top-down insistence and direction, although successful implementation often depends critically on bottom-up innovations and enthusiasms from the winners in the Darwinian struggle for survival.

GETTING STARTED: SOME DEFINITIONS AND DISTINCTIONS

Definitions

Models Versus Simulations

Discussion of M&S is complicated by terminological ambiguity. "Model" and "simulation" are often used interchangeably. In other contexts, they are distinguished, but in several different and confusing ways. Although the DMSO has issued official definitions (DMSO, 1995b), the ambiguities are long-standing and will not go away.

In this report, the panel generally uses "model" to refer to a conceptual representation of some part of the real world, perhaps something expressed in equations, diagrams, or a verbal description. Some models generate descriptions of how the system of interest or aspects thereof change over time; these are called simulation models. If such models are implemented in a computer program or human exercise, we refer to the implementations as simulation programs or, simply, simulations. In day-to-day usage, these are often referred to as models, but the panel reserves the term model for what might more fully be termed "conceptual model."

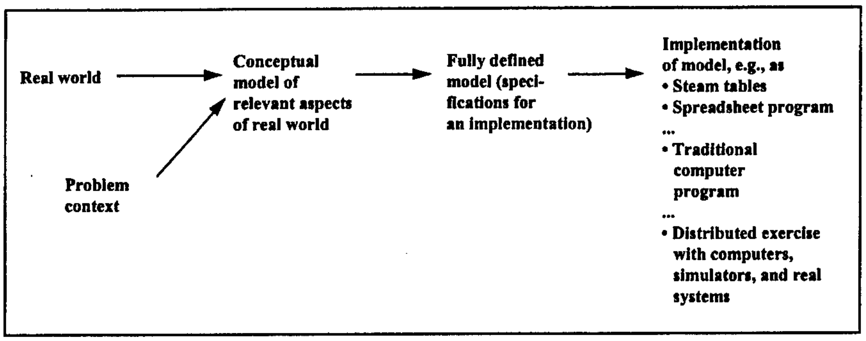

FIGURE 1.2 The real world, model, and implementation.

The phrase modeling and simulation (M&S) blurs "model" and "simulation" and, as commonly used, conveys the sense that M&S consists of computer programs or simulators such as those used to train pilots or for military exercises. However, it is important to remember that not all models are implemented as computer programs (Figure 1.2) and not all models are simulations (Figure 1.3). The first distinction is significant when talking about M&S quality, because the problems may be in the software, the ideas and designs underlying the software, or the absence of any models beyond the computer code itself. The second distinction is important because a sound approach to the problems associated with M&S should include nonsimulation models.

Having made these definitions and distinctions, the panel must now acknowledge that it also uses the term "simulation" to mean something altogether

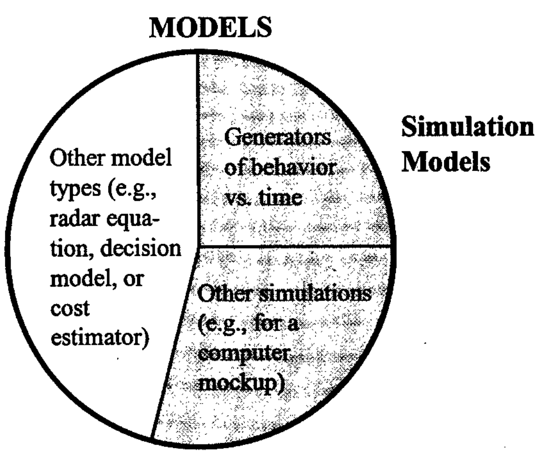

FIGURE 1.3 Types of models.

different. In particular, the panel refers to the important new activity of simulation-based acquisition. In this context, "simulation" refers to the computerized representation of something like an aircraft or ship being developed in a paperless environment—even though the representation may be purely static.[6]

Models Versus Data

Finally, there is the confusion between "model" and "data." In practice it is often not useful to emphasize the distinction, because modern well-designed M&S puts much of the content in data, providing users flexibility to change assumptions. Thus, results of a simulation are dictated at least as much by input data as by the computer program.[7] Indeed, in some cases, to change "data" is to change the form of algorithms used within the simulation. In this report, the panel treats data as part of an M&S except in a few instances where it discusses data problems per se.

Different Conceptions of Models

The next difficulty is that people have different intuitive conceptions of what models and simulations are supposed to be. At one pole of one spectrum (x-axis in Figure 1.4) are those who see models as highly flexible, constantly changing tools for use in analysis (or training). On the other pole of this spectrum are those who see models as a well-defined commodity that one should be able to procure with a warranty and use without much skill or effort.

A different spectrum (y-axis of Figure 1.4) separates those who see models as mere tools from those who see them as repositories for knowledge and mechanisms for transmitting knowledge. To the former, large complex models are an abomination because they are so difficult to comprehend and control and because they get in the way of making important points economically. To the latter, large complex models (typically implemented as simulations) embody rich depictions of important phenomena that could not readily be described in other ways: the language of mathematical equations does not go far in describing complex real-world systems transparently, much less describing or communicating their behaviors to new workers or clients.

---

[6]In other domains, "simulations" refer to such distinct things as forged documents, reenactments of historical events, play-acting of a possible negotiation, and so on. What is consistent across these diverse meanings, however, is that in all cases there is an attempt to reproduce some image, sound, or feel of a real system. By contrast, one never refers to Maxwell's equations as simulating electromagnetic phenomena. Nor is any closed-form equation defining the required characteristics of a system referred to as a simulation.

[7]That data are typically one of the primary limiting factors is discussed in Hillestad et al. (1996).

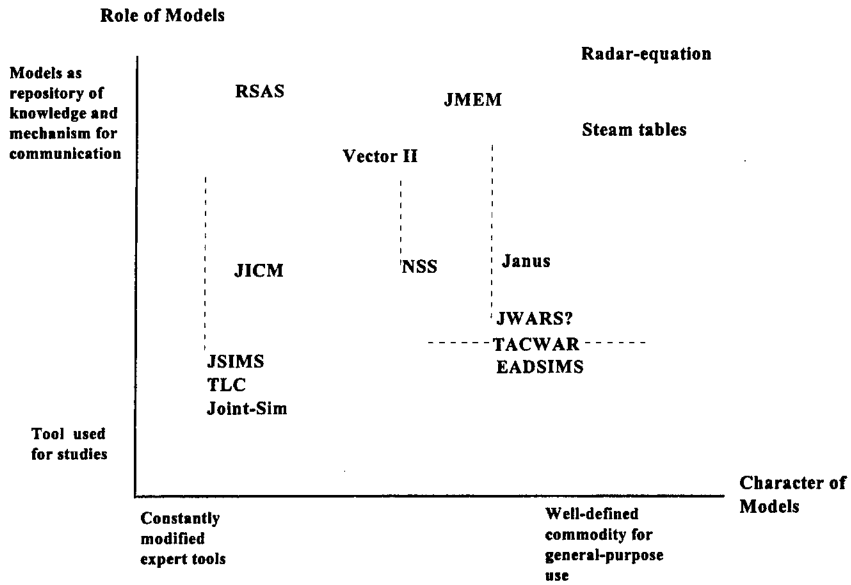

FIGURE 1.4 Different perceived roles for M&S. Dashed lines indicate ambiguity of location.

Both views have much to recommend them: they are not intended as strawmen. But given these differences, is it any wonder that there are strong and sometimes emotional tensions among individuals and groups involved in one way or another with M&S? By and large, analysts resent those who seem to see models as simple commodities (especially when the impression is given that they can be operated by relatively unskilled workers), while managers often find the views of analysts bizarre since it seems obvious to them that in fact M&S is several "systems" that need to be designed, tested, procured, and maintained and that can benefit from standards. On balance, the panel believes that virtue lies in recognizing that models serve as both tools and repositories of knowledge. This leads to conflict. However, to the extent that "big models" become less monolithic and more like environments in which to assemble tools for particular problems, the views can to some extent merge. That would be a healthy development.

Hard Versus Soft Models

Another distinction is that between "hard" and "soft" models, the former represented by, say, the engineering models used in fluid flow or sonar calculations, and the latter represented narrowly by decision models representing the adversary commander's behavior or more broadly by models dealing with individual and social behaviors. Except in Isaac Asimov's science fiction, and even then only in the aggregate over long periods of time, does anyone aspire to the

building of highly predictive social models? In the military domain, we all recognize that commanders' temperaments and life experiences matter, and that some units are much more motivated and capable than others. It is also known that special factors matter, such as deception and surprise, but also including random events such as the accidental detection or nondetection of an approaching attack such as that falling upon Pearl Harbor in December 1941.

Clearly, those involved with the different classes of models (hard versus soft) will have difficulties communicating on what can reasonably be demanded in M&S. Perversely, there has been a chronic tendency for DOD modelers and analysts to avoid representing or considering "soft factors" despite the fact that history tells us they are often dominant. This disjunction between model and reality has long undercut the credibility of most combat models with warriors, historians, and analysts willing to recognize soft factors and uncertainty.[8] It is one reason that higher-level M&S, such as campaign simulations, has seldom been of much interest to senior naval officers.

Although the methodological tensions between physical scientists and social scientists have long been recognized by interdisciplinary workers, it is less well recognized that similar tensions exist within the "hard" domain of engineering. In control theory, for example, many workers focus on near-equilibrium phenomena and on designing systems that are highly robust to large uncertainties in the environment. In this context, nonlinearities are often viewed as just another source of uncertainty. In dynamical systems work, there is a much richer view of nonlinearities and nonequilibrium phenomena, but a tendency to avoid dealing with the large uncertainties that occur in many practical engineering problems. The two viewpoints emphasize different issues and are relevant to different problems. Not surprisingly, the most exciting research directions in both control and dynamical systems involve problems with both large uncertainties and nonlinear, nonequilibrium behavior.[9]

Figure 1.4 makes the point—albeit with examples that are strongly open to question—that the various DOD M&S models have been developed with very different notions about what functions they are to serve. This makes generalized discussion of M&S difficult. Consider first models such as the radar equation and steam tables. These represent considerable scientific knowledge, but in a form that many can use readily. So also the joint munitions effectiveness manual (JMEM) records numerous specialized weapon-effect models for broad usage. By contrast, the JSIMS program is focusing on building a "tool kit" that can be used to construct appropriate training exercises. And, while it is assumed that

[8]Interested readers should consult Dupuy (1987), Davis and Blumenthal (1991), and Rowland et al. (1996).

[9]See Doyle, John, "Theory in Modeling and Simulation," unpublished, November 1996. Prepared for the panel, but based on material available at http://hot.caltech.edu and Appendix B of this report.

there will be a repository of suitable objects representing knowledge, the JSIMS program is not really organized to develop that knowledge, except for standard cases. The image conveyed is of software development, which needs databases filled out but has no particular interest in the knowledge per se. This may prove unfair, however, and current work on object models is certainly a connection to model content, so a dashed line indicates that how JSIMS ends up is yet to be determined.

In yet another contrast, mature models such as Vector II have long been seen as a repository of detailed information about forces, equipment, tactics, and terrain. The RAND Strategy Assessment System (RSAS) was also developed with knowledge acquisition strongly in mind. The JWARS program is harder to characterize. The program itself has a distinct "software" flavor, but there has to date been much less emphasis on flexibility than is visible in the JSIMS effort. Again, a dashed line is shown. Just to make another point, a dashed line is attached to TACWAR. While the original developer (the Institute for Defense Analyses) continues to modify TACWAR for particular studies and thus sees it as a tool, albeit a difficult tool to change, some of TACWAR's users appear to see it more as a fixed, configuration-controlled commodity. Rather generally, complex-model developers with whom panel members are familiar express considerable worry about misuses of their creations, which they hesitate to think of as "products" in the normal sense.

STRUCTURE OF THIS REPORT

With this background, the report proceeds as follows. Chapter 2 surveys the potential of DOD's M&S briefly, painting a very rosy future. Chapter 3 describes reasons for worry, primarily reasons related to model validity and system complexity. It concludes that a good deal of research is needed and that failure to invest adequately in such research could lead to major M&S failures. Chapter 4 elaborates on what the panel means by model quality and validity. Chapter 5 describes an important class of research that should be organized around warfare areas rather than M&S per se. Chapter 6 describes needed improvements in the conceptual, methodological, and technological infrastructure for M&S. Chapter 7 deals with challenges of assimilating and exploiting M&S technology. A collection of appendixes is intended to elaborate on and provide reference for points made in the report.

2

Technological Prospects for DOD's M&S

This chapter begins by noting the broad range of applications for M&S. Some documented examples assessing its value follow. Next discussed is the special integrative role that M&S is coming to play, which will be crucial as warfare operations become more technically and organizationally complex, and as the systems to support such operations become similarly so. The panel then offers some illustrative forecasts and visions of the future, looking both at applications (the demand-pull side of the problem) and at technology (supply-push).

APPLICATION AREAS

M&S is an enabling technology. Figure 2.1 lists some of the many applications of M&S in DOD for which M&S is increasingly essential in this role. While the applications of M&S are already numerous, the benefits of reusability and integration have by no means been realized in current systems, and cannot be until the necessary infrastructure is created.

DATA ON THE VALUE OF M&S FOR ACQUISITION, OT&E, AND TRAINING

Much of the vision discussed above is yet to be demonstrated, and it will be years before the results are in. However, there already exists a good deal of data on the value of M&S.

Acquisition	Training	Operations
• Requirements, definition, and concept exploration	• Individual	• Exercises
• Engineering design	• Small-unit	• Mission rehearsal
• Testing support	• Joint operation	• Operations planning
• Force analysis and structuring	• Joint task force	• Adaptive decision making during crisis or conflict
• (Manpower)		

Within each item are included components for combat, support, and infrastructure.

FIGURE 2.1 Partial taxonomy of applications of M&S.

Acquisition Applications: Rates of Return on M&S Investment

Table 2.1 provides data on the return on investment (ROI) on M&S investment for tools, methods, databases, and supporting techniques used to assess the lethality and vulnerability of weapon systems milestone decisions and the Cost and Operational Effectiveness Analysis (COEA) process. *The typical ROI was between $20 and $30 returned for each $1 invested* (see next-to-last column). A number of the systems are used by naval forces.

Exercise Examples

Reforger and Kernel Blitz

In the realm of exercises, one of the better known examples of using M&S was in the 1992 exercise that replaced the early Reforger exercises involving U.S. and other NATO forces. Cost savings were reported on the order of $36 million, and participants believed that training of staffs and planners was improved (Worley et al., 1996, p. 14, drawing on an earlier study by Simpson et al., 1995).

Kernel Blitz was a fleet training exercise (FLEETEX) including live ships, submarines, aircraft, and land troops. The simulation portion augmented the fleet with additional synthetic ships, submarines, aircraft, and weapons. The simulation center used several existing computer facilities (including both coasts) and existing communications capability to link to platforms. A purpose of the exercise was to show that the use of simulated assets could add realism and complexity to training exercises. It is notoriously difficult to estimate cost savings or cost avoidance due to M&S because, in practice, one could not afford to use the real aircraft, ships, and submarines included in the simulations. However, if one calculates what doing so would have cost, then the Kernel Blitz exercise saved about $16 million. Much more important, however, is that the M&S enhance-

TABLE 2.1 Return-on-Investment Data in Acquisition Work

Program	Type Analysis	Total Invest ($M)	Direct Savings ($M)	ROI	Program Result
Standard Missile SM-2 BLK IIIA	Cost Reduction	2.25	47.0	21	Accepted
Phalanx CIWS	Performance Evaluation	8.12	125.0	15	Continued
Phalanx CIWS	Product Upgrade	6.63	200.0	30	Accepted
AIM-7P Sea Sparrow	Lethality Analysis, End Game	0.7	16.0	23	Accepted
Phoenix Missile	Lethality Analysis, End Game	2.23	70.0	31	Accepted
ECM vs. AMRAAM	Lethality Analysis, End Game	0.58	10.5	18	Eval. Continues
AMRAAM	End Game	6.5	250.0	38	Continued
Bomb Fragment Data	Arena Tests	0.0825	0.9	11	Continued
BLU-109	Lethality Testing	0.0825	3.0	36	Continued
Air-to-Air Missile	Lethality plus Engagement	20.0	75.0	4	Continued
Wide Area Anti-Armor Munition	Lethality Analysis	0.75	30.0	40	Canceled
Hypervelocity Missile	Lethality Analysis	0.5	10.0	20	Canceled
ISAS	Lethality Analysis	0.75	40.0	53	Canceled
Kinetic Energy Penetrator (KEP)	Lethality Analysis	1.1	50.0	45	Canceled
JP 233 Runway Attack Munition	Lethality and Vulnerability Analysis	1.1	54.0	49	Canceled
Boosted Kinetic Energy Penetrator	Runway Vulnerability Models	2.75	130.0	47	Canceled
JAVELIN ATGM	Analytic Simulation	0.62	14.0	23	Accepted
M2 Bradley FVS	Engineering Design	0.88	30.0	34	Accepted
M1A2 Vulnerability	Damage Prediction	1.83	30.0	16	Cost Avoidance
M1A2 Block 3	Design Vulnerability	1.76	100.0	57	Terminated

SOURCE: Adapted from Worley et al. (1996), after data presented in DOD (1995), Chap. VI.

ment allowed the scope and quality of the exercise to be improved at a very low marginal cost.

A study by the Center for Naval Analyses (Neuberger and Shea, 1995) reached the following conclusion based largely on the Kernel Blitz experience: *At this point, simulation should be viewed as enriching training and increasing readiness rather than reducing costs.*

Operational Testing: F/A-18 Weapons Software Support Facility

As a third example, the panel draws on work by Michelle Bailey of the Navy's China Lake facility (see Worley et al., 1996).

The F/A-18 Weapons Software Support Facility (WSSF) at China Lake, California, is used for integration, checkout, and validation and verification of avionics software with actual avionics hardware operating as a total aircraft system. The WSSF is actually several facilities containing avionics hardware, simulations of flight dynamics, weapons simulations, and operator consoles. Several different methods have been used to estimate its cost effectiveness, but, again, the calculations are confounded by the fact that in practice one could not have flown live aircraft enough to provide the information collected in the facility's laboratory. After all, flight costs are roughly $2,800 per hour, while laboratory costs are more like $930 per hour for F/A-18s. The principal conclusion reached was that the real value added of the WSSF is that an aircraft as complex as the F/A-18 is not possible without this type of test facility. One could not fly it enough to test it. There is a danger in just looking at cost savings as the measure of whether or not one invests in M&S. As more is demanded from our warfighting systems—including the need to make them safer, more accurate, more environmentally friendly, more stealthy, longer range, and so on—one will have to demand more from our test and training systems.

M&S AS A CROSS-CUTTING FOUNDATION TECHNOLOGY

As one looks to the future, M&S will be critical not just in individual areas, but as a cross-cutting technology. To appreciate this, let us next consider "the stovepipe problem" about which so many senior leaders have railed.

Why Old Stovepipes No Longer Work

Most large organizations such as the U.S. military tend over time to break into semi-independent units with relatively little lateral communication and coordination. Such "stovepiping" (Figure 2.2) is also characteristic of the hardware, software, and M&S systems developed to serve these units. There are many reasons for stovepipes, which can be seen as modules for specialization and efficiency.

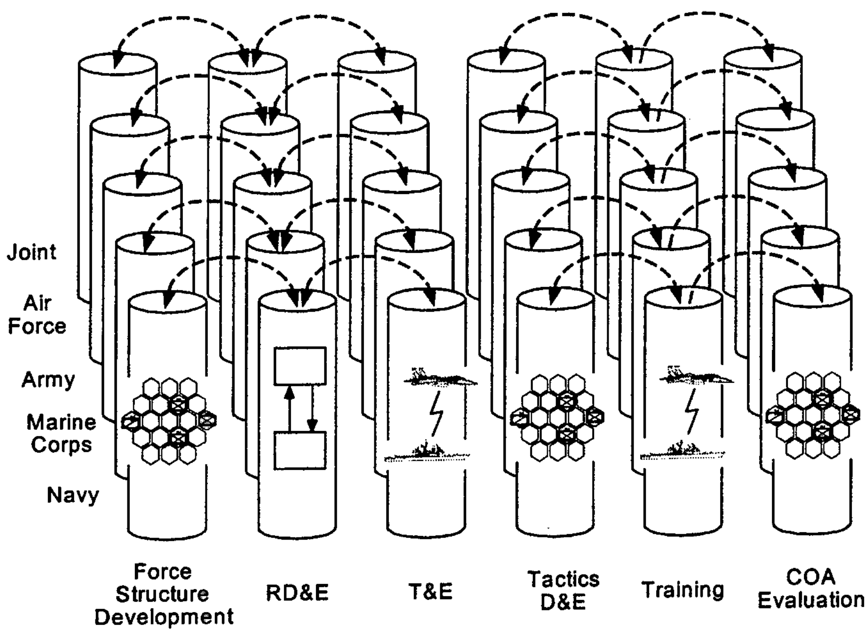

FIGURE 2.2 M&S in a joint world: Making stovepipes work.

This said, modules are supposed to connect nicely (suggested by the arrows), but the DOD's stovepipes often do not. Also, some of the traditional stovepipes are no longer the appropriate modules, as has become increasingly evident with the emphasis on jointness in technically and organizationally complex littoral operations, including precision strike with aircraft and missiles launched from ships, submarines, and air fields. These issues were noted and addressed vigorously within the Navy during the late 1980s by Admiral William Owens, who later, as Vice Chairman of the Joint Chiefs, created the Joint Requirements Oversight Council (JROC) and Joint Warfare Capabilities Assessment (JWCA) groups specifically tasked to address cross-cutting functions such as surveillance and reconnaissance, and precision strike across Service lines. Such issues are evident in *Joint Vision 2010* (Shalikashvili, 1996).

Even in the peacetime world, such stovepipes as R&D, acquisition, test and evaluation, and operations have proved troublesome as the DOD attempts to facilitate the development and fielding of advanced capabilities at much less cost and in much shorter time. So it is that we see advanced concept technology demonstrations (ACTDs) designed specifically to cut across the stovepipes and involve everyone from engineers to operators early in the acquisition process. Cross-cutting and integration are, in many respects, the name of the game.

Here M&S has a special role. To a large extent M&S will be the glue, or even the cross-pollinator. For example, it will be an essential element of the

command and control system, of operations planning, and of doctrine development. The cross-cutting will be across Service components, functional areas, and levels of command structure in each. As but one example, officers from all Services will develop an increasingly common perspective of a theater operation by training and planning with joint systems with embedded simulations. This in turn will force resolution of issues such as interservice communication protocols, a long-standing obstacle to effective operations. Further, it will facilitate standardization of planning formats and terminology.[1]

Advanced Distributed Simulation

A major component of DOD's M&S vision, probably the principal component in the view of some, is advanced distributed simulation. This now has roots extending back more than a decade, primarily to early efforts in distributed war gaming and the pioneering SIMNET program sponsored by DARPA. Much has been written about distributed simulation and the associated visions for the future, including the synthetic theater of war (STOW) concept, which is currently being pursued.[2] Two points should probably be noted here, however. First, distributed simulation is already a practical reality, something used more or less routinely by the Services and commands. Second, the cutting-edge research on the STOW concept will take many years to reach maturity because of the many technical challenges and the need to educate and train a new generation of people to assimilate and exploit the new capabilities.

Ubiquitous M&S as Infrastructure and "Cross-Pollinator"

With the diversity of application areas in mind, Figure 2.3 presents a vision for the future. The intention is to indicate that in and out of each activity such as test and evaluation will be flowing not only information, but also models and data. By no means will everything be connected to everything—whether in the sense of distributed interactive simulation or in any other way. Many workers in a given domain will spend much of their time with domain-specific tools that are never shared. But a substantial degree of reuse and sharing will occur: because it will greatly benefit those doing the work. The analogies here are perhaps best seen in today's commercial PC software, which we exploit routinely to swap manuscripts, briefings, and spreadsheets; to collaborate at intercontinental distances; and to operate in "virtual organizations."

[1]Reportedly, operations in Bosnia have been quite instructive in this regard. The United States has established an excellent command and control system with theater-level surveillance and reconnaissance. This has indeed motivated the kinds of problem-solving the panel refers to here.

[2]See "Special Issue on Distributed Interactive Simulation," *Proceedings of the IEEE*, Vol. 83, No. 8, August 1995.

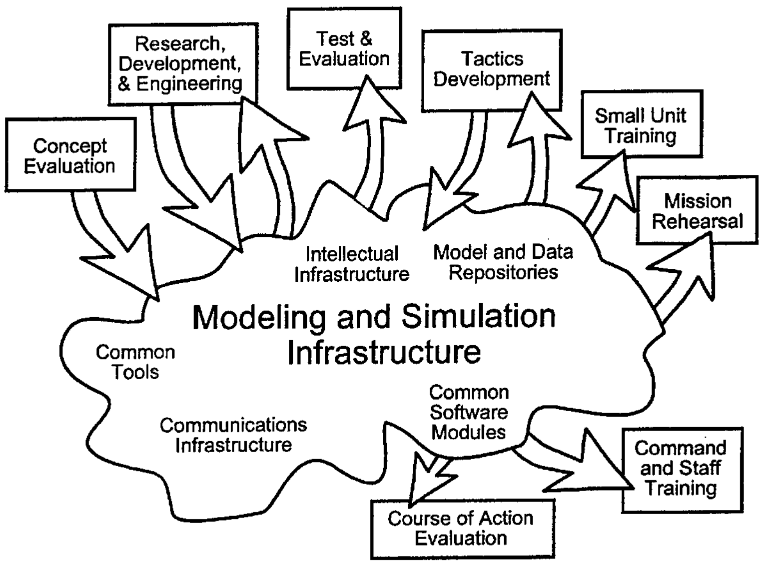

FIGURE 2.3 M&S as the infrastructure for many DOD activities.

In the future, there will be much more in this background infrastructure, which the panel discusses further later in the report. Note that a high degree of interface standardization is needed to implement such a vision. However, the infrastructure provided by M&S will be more difficult to put into place than the physical data links of communication systems and the software programs within computers. To put it differently, if achieving portability of manuscripts with formatting and graphics has been difficult and long in coming,[3] then we should expect much greater difficulties when the interfaces must communicate ideas and interpretations, not just bits. Computer scientists refer here to the difference between transmitting syntactical and semantic information, a problem familiar to commanders who learn in command-and-staff school how easily the intention of orders can be misinterpreted even if the structure of the order message is correct.

Elaborating, Figure 2.3 illustrates the notion that there is more involved than just model objects and databases. Indeed, a key element of the M&S infrastructure is commonality of intellectual constructs. To put this in perspective, readers may appreciate how universal the concepts, constructs, and notation of calculus are today, and how important they are in communication and collaboration. Similarly, fluid dynamicists worldwide can communicate readily about fluid flow. By

[3]It had certainly not been adequately achieved as this report was prepared by a virtual panel connected electronically. Compatibility problems were numerous and annoying.

BOX 2.1
Representative Benefits from a High-Quality M&S Infrastructure

- Redesign forces and doctrine for 21st century
- Enhance training, from crew member to JTF commander
- Allow commander to visualize battle at different levels of detail
- Design new systems to meet needs of all concerned, e.g., operators, designers, manufacturers, logisticians, and do so better and more cheaply
- Make optimal use of limited and expensive test assets

contrast, we do not today have commonly accepted foundational concepts, terminology, and theory for M&S. One indicator of this is the difficulty with which workers operating at different levels of resolution have communicating and cooperating. Part of the problem is technical and methodological; another part is what many see as underinvestment in and undervaluing of military science.[4] It is troubling that the words "science" and "theory" are explicitly avoided in so much discussion of DOD's M&S, apparently because of a belief that they are associated with vague abstractions rather than practical matters.[5] This belief may be understandable, but it is wrong-headed. Exploiting the potential of M&S will require breakthroughs in understanding military phenomena and in representing them mathematically and in simulation programs. To make the point more strongly, consider the contrast: excellent processing algorithms, graphics, and distributed simulation technology are available currently, but *no* major models are able to represent, for example, highly nonlinear warfare with dispersed forces and decentralized forms of command-and-control in the information age.

Returning to the theme of M&S infrastructure, a new vision of M&S is emerging in which it not only provides cross-pollination between existing, legacy stovepiped systems, but also will provide a new level of integrated support for many activities within each of the services and the previous stovepipe.

The investment in such an infrastructure would be substantial, but there are many specific benefits feasible, as suggested in Box 2.1. Consider first that the entire U.S. force structure should be redesigned for the next era of warfare. How should new force structures and doctrines be conceived and evaluated, especially in the absence of wars in which to try them out? M&S should play a major role.

Little need be said about the importance to training, because this is widely

---

[4]This is a major theme in Davis and Blumenthal (1991).

[5]One example of this avoidance can be seen in documents about the validation of models. There are references to "logical" validation and to "comparisons" with data and other models, but no reference to, say, "grounding in more fundamental theory." It is encouraging, however, that the Director of Defense Research and Engineering, and the Joint Staff now have a joint "science and technology" plan, rather than merely a technology plan.

recognized, except that the use of M&S for "training" should extend all the way from crew members to commanders. Part of the training of commanders will involve learning how to use advanced C^4ISR capabilities to "visualize" the battlefield—not only in the large view, but at several levels of detail.

New systems, of course, will be conceived, designed, and developed with heavy use of M&S. The opportunity exists through M&S to improve their suitability for all concerned—e.g., the operators, manufacturers, and logisticians—and to do so better and less expensively than in the past. It is no longer inevitable that next-generation weapon systems will always be much more expensive than the systems they replace.

The last example in Box 2.1 relates to test assets, whether they be the National Test Facility or an exercise with allies. M&S can strongly leverage the value of tests, not only for those participating in them directly, but also as a source of data for subsequent analysis.

In summary, there are many reasons for the Department of the Navy to be very interested in and concerned about the future of M&S in its domain. For this it will need a high-level policy and strategy.

SOME OBSERVATIONS, FORECASTS, AND IMAGES

A primary DOD effort in recent years has been to reform the acquisition process. The legacy process has been one of sequential activities resulting in long development times, high costs, and in some cases inability to achieve the desired product. A major problem has been separation of the various communities involved in system acquisition and use. These communities include operators, the acquisition authority, designers, manufacturers, testers, and maintainers. All must interact closely during the system development process so that the resultant product reflects all their needs. Failure to do so means, for example, that the needs of the operators and maintainers are not fully understood by the designers of the system. The result is a system that does not meet its needs or that undergoes expensive and time-consuming modification to meet them.

Figure 2.4 is an example of the vision for an improved acquisition process with emphasis in particular on the use of M&S (see Shiflett et al., 1995). In this vision, simulations are used early to better understand military needs and to test operational concepts for accomplishing missions and tasks. This permits a better statement of requirements, especially because the simulations in question bring together scientists, analysts, and warfighters (e.g., representatives from the commanders-in-chief (CINCs) and from the Service's doctrine organizations). Simulation is also used extensively for interim tests and demonstrations, again involving the ultimate users, the warfighters. The result, it is hoped, will be reduced program risk, a faster development process, and a smooth transition into the field.

In the overall vision for enhancing acquisition, the increased use of simula-

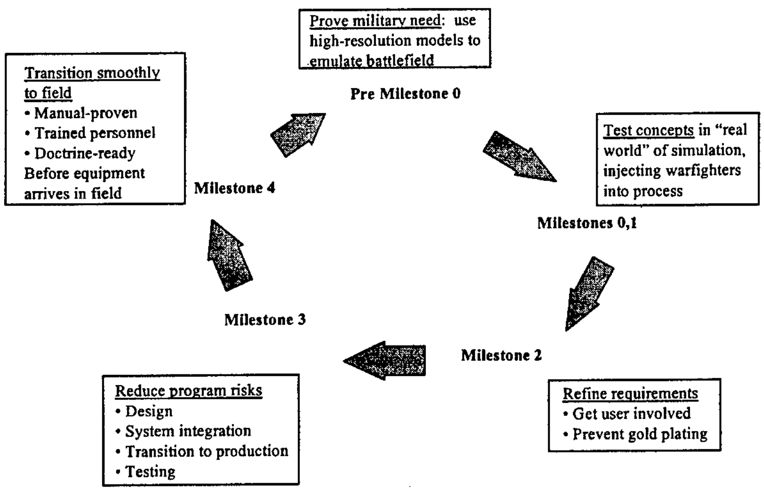

FIGURE 2.4 Simulation and reform of the acquisition process. SOURCE: Adapted from Shiflett et al. (1995).

tion is also accompanied by the widespread sharing of digital design information. The goal is an integrated design database to which all relevant parties have access. This means that the numerous engineers involved in the design process can readily share their data and will always have access to the latest design. Design inconsistencies will be reduced in this way, thereby eliminating costly and time-consuming rework in the manufactured product. In addition, design information can be reviewed by the manufacturers, who can identify design elements that would be particularly costly to produce before a commitment to production is made. The designs would be modified, and if necessary checked with the operators through the use of simulation to see that key requirements remain satisfied.

The use of simulation referred to thus far in this section is for operational purposes. This includes both virtual simulations where the behavior of the particular system (e.g., aircraft) can be examined in some detail and more aggregate combat simulations where the utility of the system at a higher mission level can be examined. In addition, engineering simulations play a key role in aiding the designers. Such simulations allow them to analyze, for example, the aerodynamic behavior and signature of an aircraft. The use of these simulations in conjunction with integrated digital design representations should allow designers to execute their tasks much more rapidly, thereby allowing a much greater set of design possibilities to be explored. The result should be a better and possibly less costly design.

Another perspective, that of investment versus time, is given in Figure 2.5.

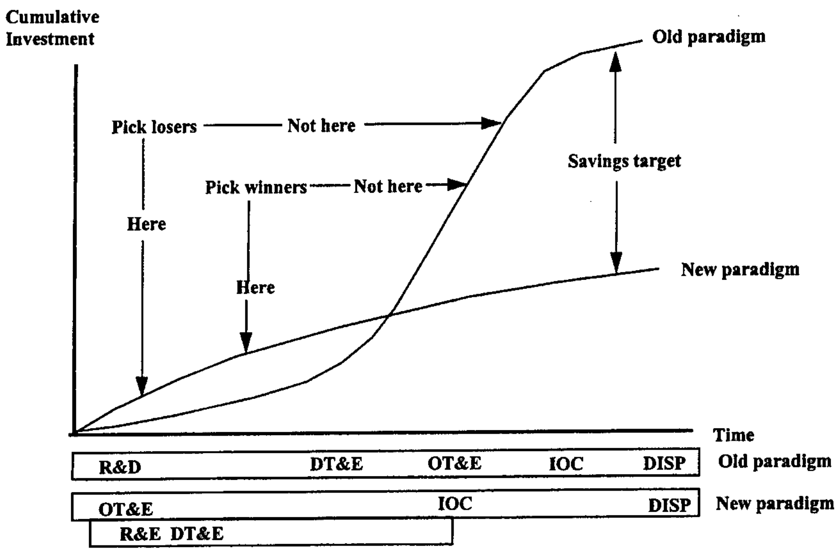

FIGURE 2.5 Two visions of investment versus time in acquisition. SOURCE: Adapted from a briefing to the panel by CDR Dennis McBride, USN, of ONR and previously of DARPA, 1996.

The point here is that by investing earlier in simulation testing of concepts, one can discard lesser designs fairly early and pick the best not long thereafter. Further, because of the heavy interaction with users and the critical use of integrated digital design representations, as discussed above, the hope is that initial operational capability (IOC) can be earlier with much lower cost. While leading to reduced overall cost, this approach, as noted in the figure, does require greater up-front cost. Program managers alone could be reluctant to make that commitment because the benefits will not be realized during their tenure. Thus, realization of this approach could well require higher-level policy direction.

The overall vision being described here is often referred to as simulation-based acquisition (see Appendix C for a more detailed discussion). In some measure, this vision is being realized now. The most prominent example of an integrated digital design representation is that used by Boeing on the 777 aircraft for describing the static configuration of the aircraft. Simulation is already widely used in the acquisition process,[6] although not in the integrated sense implied here. The integration of simulation, digital designs, and design tools has been conceptually demonstrated in the DARPA Simulation-based Design program.[7] All these examples provide opti-

[6]See Patenaude (1996). The study was conducted for the Deputy Director, Test, Systems Engineering and Evaluation, Office of the Secretary of Defense, Washington, D.C.

[7]The study can be found online at http://www.sbdhost.parl.com.

mism that the vision of simulation-based acquisition can be realized. Full realization of the vision would require significant further research and development, with key capabilities needed being multi-resolution modeling, multidisciplinary design optimization, and interface standards development (see appendixes).

At this time, three general actions appear most appropriate to move further toward the vision of simulation-based acquisition. First is the establishment of a pilot project(s) to develop simulation-based acquisition capabilities that would feed into a major naval program (e.g., aircraft or ship). This would serve to overcome the program manager's reluctance to devote his funds to increased up-front expenses and would promote demonstration, assessment, and transition of simulation-based acquisition capability in the Navy more generally. Second is experimentation involving participants from the involved communities (e.g., operators, designers, manufacturers, maintainers) aimed at developing the necessary interface standards among simulations, design data, and design tools. This would be analogous to the proto-federation experiments carried out under the direction of DMSO in development of the high-level architecture for simulation. Third is the establishment of a research program to develop those longer-term capabilities required for full realization of the simulation-based acquisition vision. This component is important since, as noted, significant research challenges still remain to achieving the vision.

TOOLS FOR DECISION SUPPORT

Sometimes in forecasting activities it is useful to develop scenarios to illustrate what might be possible in the future, with no guarantees. The purpose is to help develop potential visions. Visions can always be amended later with the benefit of more knowledge, but, in the meantime, they can contribute to innovation and communication. The panel offers the following vignette in that spirit.

Imagined Vignette— Decision Support in an Intervention Operation

The joint task force commander has a profound problem. The President has ordered U.S. forces to intervene in an ongoing war in the small nation of Blakos. The objectives include rescuing U.S. nationals and nationals of other friendly countries and securing and stabilizing events in the port-city capital of Lazune. The Secretary of Defense and Chairman, Joint Chiefs of Staff, are asking the commander for his recommended course of action. Earlier today, as his carrier battle group and a Marine expeditionary unit (MEU) sailed toward Lazune, he asked his staff to prepare alternative courses of action. They did so and presented him with two options: an assault on Lazune in approximately 6 hours with forces already available to him, or a delayed assault that would be launched in approximately 72 hours. The delayed operation would be able to employ the

82nd division ready brigade (DRB), staged from a friendly nearby island. The DRB is already en route, but would not be able to conduct an assault sooner than 72 hours because of the need to stage and to base and prepare enough appropriate aircraft. The delay would also provide the commander with a squadron of Air Force tactical fighters with substantial air-to-ground capability.

His staff has used onboard decision support systems and has drawn on expertise and analysis available from CONUS with a seamlessness that is remarkable. A portion of the staff is now developing and testing detailed versions of the courses of action with M&S that reproduces with high fidelity the time lines of all the key operational tasks that would have to be performed. At this point, they have recommended the delay option because simulation has indicated that the early assault would be defeated by Blakos forces in Lazune and arriving within the next 6 hours. Unfortunately, the delay may mean the loss of many American and friendly lives because intercepted communications indicate a Blakos intention to capture the embassies and kill their occupants.

The commander is also worried because he is not confident his staff was right in their first assessment. To be sure, their arguments seemed reasonable. However, he is troubled because models are models. He is now starting a meeting in the control center to discuss the issues in more detail. The commander thanks the staff for their work, but notes the dilemma. Is it possible that the first course of action might succeed? More specifically, he asks the staff what caused it to fail in the initial analysis. Had he asked such a question a year earlier in a similar crisis, there would have been some pained expressions because the earlier M&S had been opaque. Now, however, the staff can respond. The decision support system not only had reported the expected outcome, but also had conducted an exploratory analysis varying dozens of major assumptions. The result was a depiction of projected outcome as a function of those assumptions. To comprehend the results, it was necessary to sit in the control room with its graphical displays. There, however, he could "fly through the space of possibilities" by merely asking "what ifs?" The staff has gotten results only in the last half hour, but they are now able to report that the big problems are associated with the SA-X-25 surface-to-air missile (SAM) and the expected presence of a battalion-sized ground force in the vicinity of the embassies.

The commander now asks for more details. He learns from a sharp lieutenant that intelligence is not in fact certain whether the missiles in place are SA-X-25s or an earlier version against which U.S. countermeasures are now known to be reliable, although the simulation's algorithms and data assume otherwise. Another officer notes that the "battalion-sized force" is a motley group of poorly trained soldiers, and two of the associated companies will be straggling in for the next 12 hours. The simulations suggest that if they are only half as effective as U.S. forces would be with the equipment ascribed to them in a surprise-free battle, then they should collapse quickly under assault by the MEU's forces—if the battle group's aircraft can operate immediately without extensive counter-

SAM operations. The commander now asks whether more information can be obtained on the missiles by dispatching electronic-warfare aircraft or unmanned aerial vehicles (UAVs). As it happens, real-time reports are coming in from UAVs. Based on the radar signals they are being illuminated with, it seems that the SAMs are the older, vulnerable variety. Real-time satellite imagery is also coming in, and it indicates that the enemy ground forces are not yet taking up defensive positions and indeed appear to be disorganized. The commander also learns that there is no evidence that the Blakos forces know that his battle group is almost within striking range. Disinformation released through the news media is claiming that U.S. forces are only now heading toward the region and will not be there for a week.

Now the situation looks favorable for the immediate assault: perhaps the operation can actually accomplish its objectives, although the risks are substantial. The commander now directs a maximum-fidelity simulation, essentially a mission rehearsal, for the first option. It must be accomplished quickly because the time for decision is now. Fortunately, the decision support system has almost unlimited computer power as the result of both on-board and distributed processing. Over the next hour the commander sees the mission simulation taking form and is able to "see" it in detail. He is even able to stop it and make changes. For example, he instructs staff to make model changes to reflect the new information on the SAMs' vulnerabilities to countermeasures, and the apparent feasibility of disconnecting the SAM with information warfare attacks on the regional command post of the Blakos army, even if the command post is dispersed. Further, he is able to give contingent orders and see the simulation responding to the circumstances that arise. The most important part of the simulation involves the penetration of aircraft and their immediate destruction of two key SAM installations. Also critical is the certainty with which the battle group's long-range precision strike, from both aircraft and missiles, will be able to destroy the last two companies of enemy forces as they approach Lazune. If they will only stay on the roads, as concentrated as they currently are, it will be a duck shoot. But if they disperse or go into the jungle for a breather, the attack might be an abject failure. Success could turn into failure within minutes. The simulation has a great deal of detail available, however, regarding potential areas for the forces to rest in cover. The good news is that there will be none available for a window of time lasting about an hour. Planning continues and preparations are made in earnest for the near-term assault. Still, it may be called off if new information dictates. As we leave off, the decision support system suggests that the odds are 2:1 for an overwhelming success, 3:1 for either that or a success with losses of perhaps 100 men and 10 aircraft. The odds of mission failure, with severe losses to embassy personnel, are estimated at 1 in 10. But the odds are constantly being recalculated, and the key factors determining them highlighted for examination.

This purely hypothetical story has a number of features. For example, it postulates an M&S-based decision-support system that is intimately connected

with course-of-action development and assessment, mission rehearsal, and real-time intelligence and adaptive planning. Further, this system not only generates "expected" outcomes, but also searches and finds the variables that would change those outcomes, thereby helping the commander to focus information collection and examine some assumptions in more detail. The M&S is postulated to be not only comprehensible, but readily changeable. And it generates estimates of risk. None of this, of course, is plausible today for complex simulations.

Mid-term Tools for Commanders

The preceding material was relatively long-term speculation, but there are mid-term possibilities that are much less speculative.

Battlefield Spreadsheet

A battlefield spreadsheet (BFS) would be analogous to the financial spreadsheets now commonly used. It would be a simple model constructed by the user that will automatically propagate the effects of assumed changes in timing, forces, and so on across the battlefield. Changes in estimated number of survivors, time to move, duration of battles, and so on, would also be depicted. The first BFS may be a simple aggregate model or single entity (e.g., planning a single aircraft attack), with later extensions to variable resolution in the first case and many entities in the second.

One technique for this would be to have a powerful computer (or net of computers) playing out many runs of many variants of the scenarios in the background, with the BFS being more a display mechanism than a computation mechanism. A BFS could thus track the variability in the outcomes.

Mixed Initiative Planning

Mixed initiative planning (MIP) would be an outgrowth of the command forces/semiautomated forces (CFOR/SAFOR) technology. Currently, the semiautomated forces are fully automated at and below a certain echelon, and manual above that.[8] Work is under way to develop planning tools that work collaboratively with a planner to suggest options, check dependencies, constraints, and effectiveness, and so on. There are many tools currently available to make low-level suggestions about optimal time to launch missiles, how to find routes, and so on. But a mixed-initiative planner would do much more, for example, suggest-

[8] SAFOR systems typically provide a capability for the user to manually override the automated low-level behaviors, so the split between manual echelons and automated echelons is somewhat variable in the course of a simulation. Moreover, many scripted aggregate models would do well to achieve even this level of interaction. However, the point is that such systems do not collaborate with the user—either they run on full automatic, or they are manually steered, but there is nothing in between.

ing not only what route from A to B an armor force might take, but also what A and B should be, what position should be occupied before getting to A, what scheme of maneuver could be employed in the final assault, whether enemy forces should be fixed with artillery or maneuver forces, and so on. Such suggestions call for considerable "understanding" of the situation, tactics, and so on. The commercial analog would be the "wizards" developed by Microsoft for tools like Excel. Again, the MIP will grow out of CFOR/SAFOR and exercise-planning tools, not the commercial side.

This is quite technically feasible, and the primary obstacles are often said to be lack of management and government vision to do it.

Extrapolating these to the 20- to 30-year time span is speculative, but

- BFS and MIP will be extended to more types of weapon systems and more Joint applications of them.
- Serious modeling of precision strikes will occur, as well as continued planning for very large scale action.
- MIP will be applied in a two-sided fashion, so that plans and actions on one's own part would be countered by a simulated opponent that at least made some effort to adapt.
- MIP and BFS will be applied at multiple levels simultaneously.
- Planning tools will be extended to include realistic modeling of economic factors. This will include both the effect of combat action on the enemy's ability to fight and the enemy's effort to overcome those effects. Some work on such matters is beginning.

OTHER M&S-RELATED FORECASTS

Due to advances in many of the technologies that support M&S applications, one can anticipate all of the following:[9]

- By 2005, basic large-scale interoperability support.
- By 2010 to 2015, operationally robust support for large-scale maneuvers, including some agent-based mission-domain model checking.
- By 2015, credible simulation of combat operations before and during combat, including two-sided information warfare simulation.
- Greatly improved semiautomated forces (SAFOR).
- Speech- and natural-language interfaces to M&S.
- Agent-based mediation of input and output and of system configuration when constructing M&S for a given purpose.
- Greatly improved virtual reality systems with three dimensions and tactile and auditory stimuli. Users will enter the virtual reality and alter parameters.

[9] Abstracted from U.S. Air Force Scientific Advisory Board (1995), pp. 69ff.

3

Potential Failures and Disasters for DOD's M&S

BROAD OBSERVATIONS

Despite the dramatic success of M&S and the fact that M&S will surely "take off" in the commercial sector, the potential of M&S for the Department of the Navy and DOD may *not* be realized in the foreseeable future. Some of the DOD's most important and expensive M&S efforts may fail or—perhaps worse—end up saddling the DOD components with mediocre and inflexible tools that impede innovation and improvement of content. There is also the potential for disasters due to overdependence on M&S that appear more valid than they actually are (the Spielberg effect). This could cost lives and undercut military operations.

This chapter focuses on two causes of concern: (1) an inadequate research base and (2) the inherent complexity of emerging military systems and operations.

INTELLECTUAL AND TECHNOLOGICAL INFRASTRUCTURE

Strong Technological Base

One basic problem is that DOD's high-visibility work on M&S is dominated by content-neutral computer and software technologies such as object-oriented programming, high-performance computing, computer-aided design, and establishing the communications protocols and infrastructure for everything from collaborative, distributed, multidisciplinary, simulation-based engineering design to large-scale distributed war games. It is in these technologies that there has been

dramatic progress, which will assuredly continue because of commercial developments and DMSO efforts.

Inadequate Knowledge Base

In contrast, there has been curiously little investment in the knowledge base determining the *substantive* content and quality of much M&S—particularly higher-level M&S needed for mission- and campaign-level work. It is an open secret and a point of distress to many in the community that too much of the substantive content of such M&S has its origin in anecdote, the infamous BOGSAT (bunch of guys sitting around a table), or a narrow construction tied to stereotypical current practices of "doctrinally correct behavior." There is a need for focused research on the phenomena of combat and other military activities, both historical and prospective. This is the realm of military science.

Another shortfall in knowledge relates to theories and methods for conceiving, designing, and building models (as distinct from software). Symptoms of the problem are evident if one observes that DOD's M&S often consists of nothing more than the computer code itself: there is no separable documented "model" to be reviewed and improved, nor any way to readily understand the assumptions generating the simulation's behavior. This can hardly be a comfortable basis for decision support.

COMPLEX SYSTEMS AND THE NEED FOR HUMILITY

The inherent complexity of the systems and force operations that DOD is attempting to simulate introduces new difficulties (Appendix B). Throughout its efforts, the panel was concerned by the degree to which many forecasts are extrapolating unreasonably from the Boeing 777 experience and from M&S successes in lower-echelon training to imagined systems of extraordinary complexity. Generally speaking, the types of complexity being considered break into three distinct, but not wholly unrelated cases:

- *Localized systems.* These are systems that are highly complex, but are designed from inception as one system. Examples could be an aircraft or a ship, as well as a very large scale integrated circuit.
- *Systems of systems.* These are distributed (typically information) systems for which the final overall configuration is not known during design of the individual components. In fact, overall system configuration often varies depending on the particular application or circumstances.
- *Combat operations.* Here the complexity is due to the interaction of multiple combat systems and forces with one another and with the physical environment in which they are operating.

Significant increases in complexity are occurring in each of these cases. For example, the software in modern weapon systems is perhaps 100 to 1,000 times larger than in weapon systems one or two generations removed, now often totaling over 1 million lines of code. Distributed network systems such as the Internet have become so extensive that the causes of network outages often take hours to ascertain. While anticipated combat operations have become smaller than those in cold war scenarios, the operations themselves have become more complex in the sense that far greater quantities of information are being exchanged and interaction between units can be much more coordinated (e.g., in combining the effects of a small number of dispersed forces to have a large cumulative effect).

This increased complexity means that it will be harder to design systems and to predict their behavior. Some might argue that in the engineering domain, modern design tools should overcome these difficulties. While there certainly have been impressive advances (e.g., in computer-aided design), there is a feeling with at least some members of the engineering community interviewed in the course of the study that they have nearly reached the limits of complexity that can be addressed with current tools and methods. Challenging examples that could lie beyond current approaches include future generation networks and very large scale integrated circuits. Indications of the difficulty of building complex engineering systems are given by some of the well-known "disasters"—explosion of the Ariane missile, inability to build a next-generation air-traffic control system, outages in telephone and power systems, and even the problems with the baggage handling system at the Denver airport.

What this means from the perspective of M&S is that, if the systems are harder to design and build, then it certainly will be harder to model and simulate them. But, on the other hand, this makes the need for M&S all the greater in designing systems and predicting their behavior. Some of the factors that pose particular problems for the M&S of complex systems are as follows:

- *Scale.* As the number of elements in a system increases, or as the dimension or resolution of the space in which it is considered increases, the time to perform relevant calculations (e.g., seek optimal behaviors) can increase as a polynomial or even exponential of the number of these elements.
- *Nonlinearity.* A given component in a system influences a second, which influences a third, which interacts back on the first, and so forth. This behavior obviously becomes highly complicated as the number of components increases. Such interaction can lead to nonintuitive results.
- *Heterogeneity.* The different components are of a fundamentally different nature or the modeling techniques that must be brought to bear to fully understand a system involve different disciplines (e.g., structural mechanics, aerodynamics, human factors, and propulsion in the study of aircraft performance). This means establishing interrelationships among the disciplines, which can be complicated by inconsistent formalisms and assumptions among them.

New approaches in M&S will most likely be required to address these factors. Two particular points are of note. First is the matter of interaction. For systems with large numbers of components, often the sheer nature of the interaction between the components becomes the dominant factor, rather than the detailed properties of the components. An example is given by the transportation simulation (TRANSIMS) developed at Los Alamos National Laboratory. Realistic, large-scale automobile traffic behavior is predicted by using an approach that describes driver behavior with only simple interaction rules. This result is perhaps counterintuitive to traditional approaches that would have focused on describing individual entities (the automobiles and drivers) in great detail to get realistic traffic behavior. The general implication is that high-resolution description of components may not always be the appropriate way to model the behavior of complex systems involving a great many components. Rather, a formalism must be developed that gives equal (and perhaps greater) weight to describing interaction as it does to describing the properties of the individual entities.[1]

The second matter is that of uncertainty. No matter how careful one is in modeling a system, there always will be uncertainties. Within single disciplines, extensive effort has been devoted to characterizing and understanding the implications of uncertainties. The matter becomes particularly acute when multiple disciplines are involved, where the results of a model from one discipline are used in conjunction with those from another discipline. The issue is that there is not, in general, an understanding of how the uncertainties from the first model will influence the predictions of the second model, and so forth as even more disciplines are involved. This propagation of uncertainties could lead to highly erroneous results. What is needed is a formalism, preferably a domain-independent one, that would allow the characterization of the propagation of uncertainties.

In an important sense, one can say that complexity and uncertainty are closely related. Namely, it is the desire to reduce uncertainty that can lead to complexity. For example, the complexity of a modern air-to-ground attack missile, in terms of the number and interaction of its sensing and guidance components, results from the need to reduce the uncertainty in the location of the target that is being attacked. Likewise, the complexity of a large-scale intelligence system, in terms of its data collection and processing components, results from the need to reduce uncertainty about the enemy. Many engineering disciplines take this perspective—control theorists construct feedback mechanisms to cope with uncertainties in the performance of actual systems, and statisticians devise sampling schemes to reduce uncertainties in measured properties. Given the fact that treatment of uncertainty is central to many individual disciplines, the hope is that a more

[1]This perspective could have profound implications for combat modeling, where the tendency often is to want to go to greater and greater levels of resolution of the entities. Preliminary explorations of this perspective are given in Ilachinski (1996a,b).

comprehensive and powerful approach to treating it can be established by having these disciplines collaborate in addressing it.

As the situation now stands, these matters of interaction and uncertainty are daunting challenges. But, as noted, the difficulty of designing and predicting the performance of complex systems makes the need for M&S of such systems all the greater.

4

Dealing with and Improving DOD's M&S

SIGNIFICANCE OF THE ISSUES

As discussed in Chapter 3, the quality, substantive content, and proper use of models and simulations are significant issues in being able to achieve the full promise of M&S. These were indeed a motivation for the panel's original tasking, as noted in Chapter 1. We now turn to those issues.

Models as Repositories of Knowledge

One reason that DOD's models, especially the higher-level models, have significant validity problems is that many model builders treat them merely as tools to be manipulated as required in the context of a particular study. With this view, a developer may well set much lower standards for model quality than would have been set if the developer believed that these models were products that would be handed over to others for their own use. From this "tool" perspective, model and simulation quality is associated with the study rather than with the models themselves. Another consideration here is that the historical experience has been that attempts to build comprehensive general-purpose models have often collapsed under their own weight.

The "model as tool" perspective has immediate practical value in that it permits the particular study to be completed more efficiently; however, it has very bad longer-run consequences. These untoward long-run effects arise in part because (1) they defeat the potentially positive role that can be played by M&S and (2) they defeat the longer-run benefits that can come from good software process. More specifically,

- Simulation models are becoming unique repositories of knowledge about complex systems. As such, models must be carefully specified, fully documented, and designed to be evolvable as knowledge about the system or phenomenon being modeled is gained.
- Simulation models are increasingly becoming an important mechanism by which knowledge is communicated and passed on.
- In the future, simulation exercises will become a major vehicle through which the intuition and insight of military officers about combat will be developed and enriched. Consequently, it is of great importance that those models convey reality as much as possible.
- Well-developed, well-documented, and well-calibrated models and simulations can be used in many studies. When such models are reused, a wide audience receives training on them, improvements to the models can be suggested, and an evolutionary process can be established through which those models can be continually improved.

Modeling and simulation is already playing the role of being a repository of insight. For example, when a new analyst begins work in an organization, it is often the case that his or her education is centered on "learning the model." The organization's principal model is the frame of reference for discussion and tasking. Even though the analyst may be told the aspects in the model are realistic and unrealistic, the model as a whole is frequently at the core of his or her work. In such a case, it is important to institute an improvement process within which the model can continually evolve and be upgraded.

Models, then, are far more than tools. Appropriate models can represent and communicate our knowledge. Inappropriate models (or use of models) can distort situation assessment and choices of alternative courses of action, whether it be in the choice of weapons systems or the choice of operational strategies and tactics in the midst of war. Successful military operations are increasingly dependent upon the use of sound, well-documented models. Moreover, the advent of distributed simulation and the evolving character of DOD command and control systems serve as additional reasons why users of models will often not be part of the same organization that developed them. These trends put a far greater onus on the model or simulation developer, namely, in distributed-simulation applications (and in the future world of M&S in which frequent use is made of repositories), models and simulations must be increasingly well developed and well documented and must be constructed so that they can be reused across studies and improved as the needs arise.

The potential of M&S for impact on the Department of the Navy, both good or bad, will increase greatly over the next 30 years. The Department of the Navy should have a great interest in capturing the potential benefits while avoiding the many potential problems, but this will not happen without focused institutional attention and some significant changes in current policies toward M&S. For

example, the issues of model quality and content have gotten short shrift for years relative to the attention paid to the more tangible underlying computer technologies such as graphical interfaces, processing power, and network connectivity. Model quality and content need much greater institutional attention.

THE MULTIFACETED NATURE OF MODEL QUALITY

What determines the "quality" of a model (including simulation models)? Perhaps the most important elements of quality are the following:[1]

1. *Knowledge.* Validity of the knowledge represented.
2. *Design.* The structure used to represent the system, which can affect the model's clarity and appropriateness, as well as its maintainability.
3. *Implementation.* The faithfulness and soundness of the model's implementation as a computer program.
4. *Software quality.* The quality of that program as software, taking into account considerations such as comprehensibility, modifiability, and reusability.
5. *Appropriateness of use.* The way in which the model is used substantively, particularly with respect to dealing with uncertainties. Assuming that programming is correct, it is often difficult to assess a model's quality outside of context.

In what follows, the panel discusses these issues in two pieces. First, the issues of knowledge, design, and use will be considered (items 1, 2, and 5) because they are often closely intertwined. In doing so the panel will use the theme of model uncertainty as a focal point. This is unorthodox, but quite useful for the present purposes. It will also help motivate the approach the panel suggests to research, an approach that recommends both research focused on warfare areas (Chapter 5) and more fundamental research in modeling theory (Chapter 6).

UNCERTAINTY AS A CORE REALITY IN BUILDING AND USING MODELS

First, of course, there is uncertainty about how models should repesent the world. That is, there are uncertainties in our knowledge. However, going beyond this, a central reality is that models and data will remain inherently imperfect—especially with respect to higher-level matters such as those depicted in operational- and campaign-level M&S. Therefore, the Department of the Navy must

[1]A different breakdown might be in terms of knowledge, quality as a software artifact, and both cost and benefit to users. See also the discussion in Appendix B.

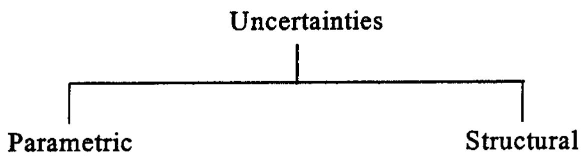

FIGURE 4.1 Taxonomy of model content uncertainty problems.

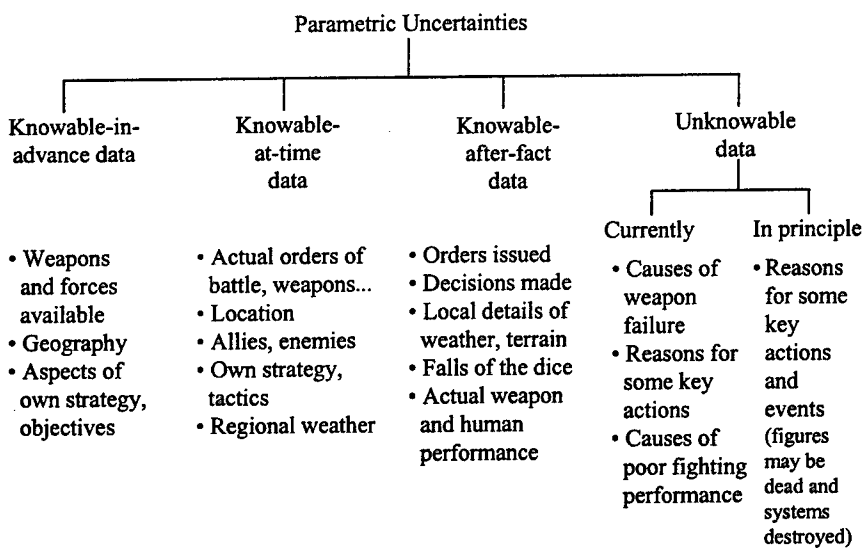

FIGURE 4.2 Taxonomy of uncertainties in parameters.

also improve its willingness and ability to deal with uncertainty in models and with the predictions made by them. This will require a major cultural shift (throughout the DOD community).

An initial taxonomy for dealing with uncertainty in models is presented in Figure 4.1. Uncertainties can lie in model parameters and, more fundamentally, in their basic structures.[2]

Figure 4.2 provides a taxonomy of parametric data. Starting on the left, we see the class of data that are knowable in advance, at least with enough effort. This includes data on friendly and enemy weapons and forces.

There is then a class of data that are unknowable at the time of force-planning decisions, but knowable at the time of an actual contingency. This

[2]A different breakdown involves limitations of understanding, computation, and measurement.

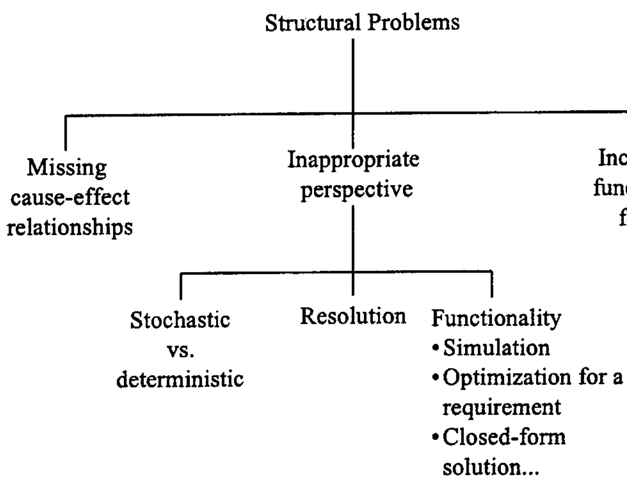

FIGURE 4.3 Taxonomy of structural problems in models.

includes, for example, the real order of battle for both sides, and whether the United States and the enemy have allies.

The next class is data that are not knowable until after the battle or war (and perhaps not even then as a practical matter). For example, the simulated battle may depend on particular decisions made by the commanders of both sides. One might estimate what those decisions would be by drawing upon logic, doctrine, and even cognitive modeling of individual commanders if one knew who they were, but the results of such efforts would still be estimates. The real decisions would not be known until afterward. Similarly, the actual cloud cover over a particular bridge at a particular time is known only when the time comes. But such things can be determined afterward from records. Thus, models using such parameters are by no means unscientific or circular. They are, however, limited in their ability to predict.

Still other classes consist of data that are unknowable. That is, there is some fundamental parametric uncertainty, even after the fact. For example, the morale of a key unit might be bad at a critical time because of illnesses, exhaustion, poor leadership, or a host of other reasons. The unit might be destroyed, along with its memories. Some such unknowable data can be represented as randomness, but it may be randomness of a complex sort.

In summary, even a "perfect" model may have limited predictive capability. It may be quite useful for description and for developing insights about what *might* happen. And it might be predictive in some circumstances where the unknown factors are not very important.

Parametric uncertainty, then, is quite large. However, it is only part of the story. Figure 4.3 provides one taxonomy of structural problems in models, that

BOX 4.1
Examples of Structural Problems

- Ignoring consequences of ineffective ASW
- Ignoring effects of warning time on ship survival
- Overestimating stealth effectiveness in at-sea operations
- Not reflecting effects of logistics flow ashore on maneuver warfare
- Assuming independent events in mine-effectiveness calculations
- Assuming equally effective friendly and enemy personnel
- Not representing battle damage implications of small crews
- Assuming static, nonadaptive tactics
- Ignoring countermeasures to long-range precision strike
- Assuming linearity and independent events in evaluating PGM versus moving targets

Order of magnitude errors possible, also "tail" effects

is, problems embedded in the choice of entities and, within processes, the logic and algorithms. Since we often do not know precisely what the "correct" structure is, these problems correspond to a different form of uncertainty.[3]

To make this less abstract, Box 4.1 lists some examples of structural problems. Some of them are probably quite familiar to military readers; other are known primarily to specialists. Commenting on the last two items as examples, the panel notes that much modern planning for next-generation warfare depends sensitively on the effectiveness of long-range precision strike. Theater-level analyses, however, typically estimate that effectiveness by merely concatenating planning factors about kills per sortie (or volley) and sorties (or volleys) per day. Such analysis is a crude linear and deterministic approximation that can be wrong. To a commander concerned about troops in the field, troops allegedly protected by long-range fire, it might be of interest to know that the correct mathematics would involve a probability distribution for the effectiveness of that fire, and that—even if the "best estimate" indicated that the enemy forces would be destroyed before engaging small friendly units—there might be a substantial "tail," that is, a substantial probability that many enemy forces would in fact penetrate and engage.[4]

[3]See also Appendix B on virtual engineering, which gives a related but somewhat different taxonomy.

[4]See also the discussion in the chapter on analysis and modeling in Defense Science Board (1996a), Vol. 2. See also Appendix J on probabilistic dependencies.

The overall point of Box 4.1 is to dispel the myth that results depend only on model data (i.e., the assumptions about parameter values). In fact, they also depend on built-in features.

APPROACHES TO DEALING WITH UNCERTAINTY

Coping with ubiquitous uncertainty, both parametric and structural, requires something increasingly referred to as "exploratory analysis," as distinct from analysis focused on the implications of allegedly best-estimate assumptions and some modest sensitivities. Described more fully in Chapter 6 and Appendix D, exploratory analysis is only now becoming computationally feasible as the result of massive increases in computer power.[5] However, current M&S has not been designed for this kind of analysis under uncertainty. Nor have civilian and military leaders, or analysts for that matter, been educated to approach problems with a full confrontation of uncertainty. Changing this circumstance for next-generation M&S is therefore important. This will require basic research, education, and cultural changes.

"DOING BETTER" ON MODEL CONTENT: NEED FOR MANAGERIAL CHANGES, NOT JUST TOKEN EXHORTATION

Given this quick review of types of problems, what can be done? The most important point is that doing better requires commitment to model content. The panel suggests an approach for the Department of the Navy as indicated in Box 4.2. It starts with an expression of top-level interest and concern. It then involves a strategy. The principal notion is that model quality tends to be highest when the modeling has been done by people working on relatively specific problems. Thus, the strategy does not call for increasing money for M&S content per se, much less for investing in "hobby shops," but rather for investing in research in each of a number of important warfare areas that need such research to avoid serious errors that might cost lives or lose wars. For each such program the panel suggests providing terms of reference that explicitly demand empirical and theoretical research, not mere modeling. Further, the panel recommends review by science advisory panels, whose members are highly qualified in terms of education and research experience.

The panel also sees the need, therefore, to increase the supply of military officers qualified to oversee the M&S efforts. The qualifications would involve

[5]See Bankes (1993, 1996) for a technologist's perspective. See Davis et al. (1996) for applications to force-planning analysis.

BOX 4.2
An Approach to Improving Model Content (Quality)

Possible Instruments and Approach

- Policy
- Explicitly recognize problems of model content
- Highlight need for research on knowledge base

Strategy: Establishing Good Exemplar Programs in Key Warfare Areas

- Processes that include
  —TORs spelling out research requirements for each such program
  —Review by science advisory panel
- Qualification standards for managers of M&S
  —Degrees
  —Experience
- Renewed support for top-quality long-term level-of-effort research
  —Promotion criteria for M&S managers and agencies
  —Long-term research and VV&A
  —Open "publication" and peer-review debate

solid technical education (e.g., M.S.s or Ph.D.s requiring research), experience, and, perhaps, certification in a "short course" for technically educated mid-career officers about to take on managerial responsibilities involving M&S. There is already a shortage of such officers, with the result that some officers find themselves managing M&S with little or no relevant background except for general-purpose managerial skills.[6]

The panel also sees need to improve incentives for young officers to work on model quality, and to do a better job of matching background to responsibility. More investment in education, and changes in promotion criteria, are needed.

[6]This problem cuts across the services, but with respect to the Navy, it is probably worsened by the lack of a substantial and prestigious analytic shop analogous to Air Force Studies and Analysis in its peak years or to the OP-96 organization disbanded in the 1980s. The panel notes, however, that it is not advocating that M&S and analysis be the province solely of "analysts." Involving military officers with substantial operational experience, and forcing modelers and analysts to become acquainted with operational realities, is critical. It is now arguably more feasible with the advent of distributed simulation as a key factor in training and exercises. See Davis (1995b).

VERIFICATION, VALIDATION, AND ACCREDITATION

General Observations

Any report discussing model quality, content, and validity must address what DOD calls verification, validation, and accreditation (VV&A). Much has been written about the subject, however, so the panel merely touches here upon highlights relevant to the current study.[7]

Roughly speaking, verification testing establishes whether a computer program correctly implements what was intended by the designer. A verified program should run without crashing, should accomplish its numerical calculations correctly, and so on. In practice, verification testing not only uncovers a variety of "bugs" ranging from typographical errors to incorrect bounds on variable values, it also uncovers errors or shortcomings that trace back to design (e.g., omitted logical cases). Nonetheless, verification's purpose is primarily to test implementation, not the correctness of the underlying model. Validation is the continuing process of establishing the degree to which the model and program describe the real world adequately for the intended purpose. Accreditation is an official determination that a model is adequate for the intended purpose.

In the early 1990s, when DOD established the Defense Modeling and Simulation Office (DMSO), one priority was to establish procedures for assuring validity. DMSO prepared a formal instruction to DOD components, defining terms and objectives and requiring that the components develop VV&A plans. The DMSO also sponsored or encouraged a number of efforts to define VV&A issues with some care and establish guidelines that could be adopted by project leaders and managers throughout the DOD. There now exists a good deal of related documentation.[8] One prominent feature of that documentation is a community consensus on a number of important principles, notably,[9]

1. There is no such thing as an absolutely valid model.
2. VV&A should be an integral part of the entire M&S life cycle (i.e., quality cannot be "inspected in," to quote a well-known aphorism).
3. A well-formulated problem is essential to the acceptability and accreditation of M&S results.

---

[7]For a range of current views of VV&A, see Sikora and Williams (1997), Muessig (1997), Chew (1997), Stanley (1997), Youngblood (1997), and Lewis (1997).

[8]The most comprehensive DOD study on the matter may be Defense Modeling and Simulation Office (DMSO, 1996a) which describes a number of consensus principles, prescriptive material drawn from academic and industrial experience with model development and testing, and illustrative formats for reporting results of VV&A activities. See also Davis (1992), MORS (1992), Youngblood et al. (1993), and Sanders and Miller (1995).

[9]Many of these principles are discussed by Hillestad et al. (1996) based on experiences at RAND in developing and maintaining campaign models.

4. Credibility can be claimed only for the intended use of the model or simulation and for the prescribed conditions under which it has been tested.
5. M&S validation does not guarantee the credibility and acceptability of analytical results derived from the use of simulation.
6. V&V of each submodel or federate does not imply overall simulation or federation credibility and vice versa.
7. Accreditation is not a binary choice.
8. VV&A is both an art and a science, requiring creativity and insight.
9. The success of any VV&A effort is directly affected by the analyst.
10. VV&A must be planned and documented.
11. VV&A requires some level of independence to minimize the effects of developer bias.
12. Successful VV&A requires data that have been verified, validated, and certified.

These principles are significant in large part because they are not obvious to many users of M&S, and even to many military officers and civilian officials who find themselves involved in M&S. Indeed, it is apparently natural for many individuals to assume that models can be tested, once and for all, and either certified or rejected. Such people are implicitly seeing models as commodities (or as physics models).

Implications for Management

The consensus principles accepted by workers, if not their managers, have many implications for model management (Davis, 1992). Perhaps the most important is recognizing that model quality must be built in from the outset and that assuring high-quality knowledge content will often require *years* of continued effort and a permanent effort to maintain and update. With respect to VV&A,

- VV&A should be seen as merely one part of a much larger effort to ensure model quality, an effort that depends for its success on understanding the relevant phenomena, representing it in modeling terms appropriate for intended applications, implementing that representation in a computer model, and seeking out aggressively the data needed to use the model.
- It is not generally possible to assess a model's validity once and for all. First, validity depends on contextual details; second, most models and/or their data change frequently. Indeed, the best models for a particular effort—whether training, acquisition, or operations related—are often assembled (or patched together, to use an older metaphor) for that specific purpose.
- An enlightened approach to model development will anticipate and support long-term continuing research on the warfare phenomena at issue, as well as

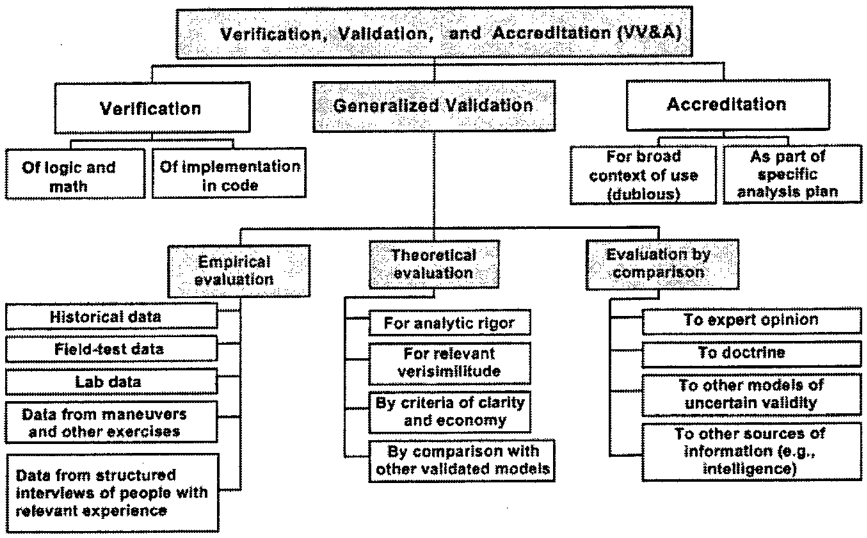

FIGURE 4.4 A taxonomy of VV&A methods to assist managers in mid- and long-term planning to ensure model quality.

related data collection. It will also support a broad range of information collection and empirical testing.

Consistent with this, managers and sponsors of M&S need to support the knowledge base with a broad range of information collection and empirical testing, as suggested by the shaded portion of Figure 4.4. This suggests the need for a long-term *program* to do so, because, in practice, shorter-term efforts such as the effort to do the computer coding for a model seldom if ever can do more than a tiny fraction of what Figure 4.4 suggests. In organizational terms, Figure 4.4 implies that the research may need to be supported by an office with a longer time horizon and a research orientation, even though the intention is to have the research be strongly focused. Over a period of years, such an organization could build a strong knowledge base in a given area with the combination of activities suggested. Note that these include not only the commonly used evaluation by comparison with expert opinion, doctrine, and other models, but also empirical work and theoretical studies.[10]

[10]Our discussion is too brief with respect to verification. Verification testing is essential and often insufficiently pursued. Even numerical methods are sometimes inadequate to deal with nonlinearities. A key to success is assuring quality up front with clear and well-reviewed designs and routine component testing along the way. Also, many software engineering tools can help a great deal if employed from the outset.

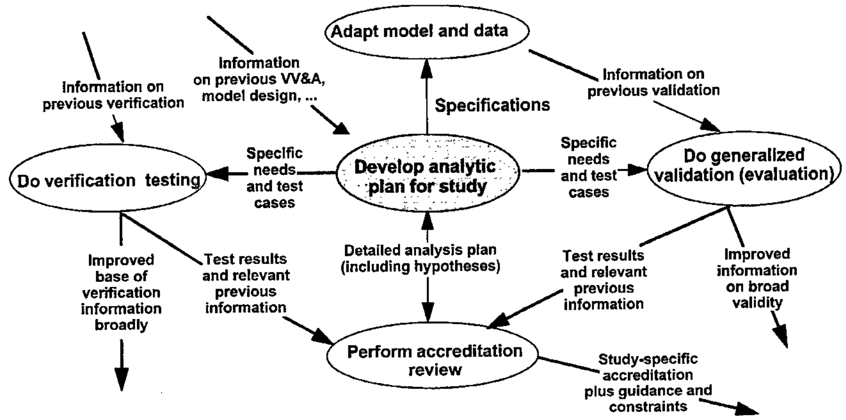

FIGURE 4.5 VV&A as elements of an application-centered process over time.

A second managerial/organizational implication of the conclusions about VV&A principles is that accreditation should be seen as applying to a specific activity such as a particular study or exercise, and not to a model. The panel recognizes that this runs counter to the natural desire of the military and civilian managers who want to have their models certified once and for all, but the conclusion is central. If one truly cares about M&S quality, rather than about having gotten a "check mark" for that M&S, then the context of use is critical. To be sure, an M&S can and should be reviewed and certified for basic soundness and performance, and most important, perhaps, for whether it is adequately documented with "truth in advertising" about what its strengths and limitations are, and cautions for users. Such reviews should be strongly encouraged. However, they have essentially nothing to do with whether a subsequent application is sound. That application, for example, will be likely to require a special database. The data will probably dictate results. What significance, then, would an earlier model review have? Nor should the data be held constant, because in practice that can often undercut the quality or relevance of the work.

The inexorable conclusion, again, is that the quality of an application must be accredited with full appreciation of the specific context. This, then, is not so much a VV&A activity for an M&S as it is a more general review of a substantive activity such as a study or fleet exercise.

If one takes this view, then Figure 4.5 suggests how to implement it. It shows the critical element of accreditation as the development of an analytic plan for the study (or exercise, etc.). This plan will then dictate model adaptations, specialized verification and validation testing, and the criteria to be used in that testing. The result (bottom right) is a study-specific accreditation, probably with guidance such as "Do not purport to reach conclusions on . . . because the study

is not adequate for that purpose. Further, report results as ranges over the specified uncertainty band, because 'point results' would be misleading."

As of the time this report was being prepared, the Department of the Navy did not appear to have a managerial concept for validation-related research. Further, its VV&A plan was deliberately permissive, in keeping with the Navy's tradition of decentralized activity on M&S. That approach may be desirable in many respects, but the failure to plan supportive research activities is a problem—albeit, one DOD-wide in its scope.

An Alternative to Emphasis on VV&A Processes

While VV&A is unquestionably important, it is doubtful that a focus on bureaucratic process will greatly improve the quality of models and simulations. Too many of the problems start at the outset as noted above. In Chapter 7 the panel recommends a "market-oriented approach" that emphasizes increasing the testability of M&S and then exposing the M&S to extensive "beta testing" by organizations such as the Naval War College. The panel also recommends demanding and then exposing to outside scientific review the "conceptual models" on which M&S should be based. Today, such conceptual models often do not even exist and hence cannot be reviewed, but that situation should change. The current emphasis on building and publishing object models is an important step in the right direction, as is work on common models of the mission space (CMMS).

5

Focusing Warfare Research and Improving M&S

BACKGROUND

Given the need for research to improve the knowledge base on which M&S is based, how best might the Department of the Navy (and DOD) go about it? Although the issue is often posed as building better M&S, that is arguably an instance of the tail wagging the dog. Do we ask aeronautical engineers to build better models or to build better aircraft? Do we ask economists to build better models or to clarify important issues such as how to define a cost-of-living inflator? Why, then, do models and simulations have such a prominent place in current DOD thinking?

In fact, there are several reasons. First, M&S products (as distinct from constantly changing personal tools) are needed to achieve the objectives of distributed training, exercising, and planning. Second, M&S products are needed to achieve the improvements in effectiveness and efficiency associated with making reusable objects in a generally available repository. The third reason, however, is one mentioned early in the report: that many people, particularly managers and software technologists, *think of* models as commodities and do not worry particularly about where the knowledge comes from to support the models.[1]

The point here is not to criticize that view, because it has its place. Indeed, the insights from software engineering and management have great value for military science in which models and simulations are important, which includes the study of most complex military operations and phenomena. However, there

[1]Also, many "modelers" are more focused on constructing interesting programs than on applications. This violates principles of what operations researchers are taught, but it is a sociological fact.

is a need to rebalance the situation by emphasizing that where the objective is to understand the nature of war and other military operations—especially the nature of future war in the information era—the focus should be on research rather than model building per se.

Research Versus Simulation Building

To appreciate the significance of research versus simulation building, consider what often happens when people define the objective as building an M&S. A committee goes about constructing wish lists, which are later translated into an expression of requirements. A request for proposals (RFP) is then issued or the tasking assigned to a government laboratory or federally funded research and development center (FFRDC). A contract is let and work proceeds. But the work is typically construed to be *building software*. The team might have, for example, a chief modeler/designer, a software designer, several modeler/programmers, and some specialists in graphics, databases, and operating systems. And, because the model will need data, there may be one or several individuals, as well as representatives from sponsors, actively involved in building databases (e.g., for orders of battle, temperature profiles in different portions of the ocean, weapon effectiveness, sortie rates, and so on).

Now, all of this may sound reasonable and industrious, but it is quite different from what would happen if the objective were seen as understanding the subject area, with a model as a possible by-product. In this case, the team might include scientists, engineers, operations researchers at least as interested in phenomenology and conceptual models as programming, historians, and psychologists (e.g., for interviewing experts)—as well as operations-experienced military officers. Results might include learned papers on various aspects of the phenomenology and other papers discussing future doctrinal options. Unfortunately, there might *not* be any products directly usable by builders of M&S. There might be no rigorous models at all, or they might not "fit" well into the larger scheme of things.[2]

Ultimately, what seems to be needed is a synthesis (Figure 5.1). There is a need for research, but that research could be accomplished with the recognition that it will be used to feed the building of M&S, and it could be accomplished with the same common model of the mission space (CMMS) as used by the M&S builders. Further, the M&S designers could base their designs on concepts emerging from the research rather than imposing their own concepts. The result would be an M&S better able to accommodate future research results as well, rather than

[2]As an example of this, there have been a number of interesting historical studies on when and how battles are won and lost, but they have seldom related easily to the simulation models on which DOD depends. Incorporating their insights, much less their data, has been difficult.

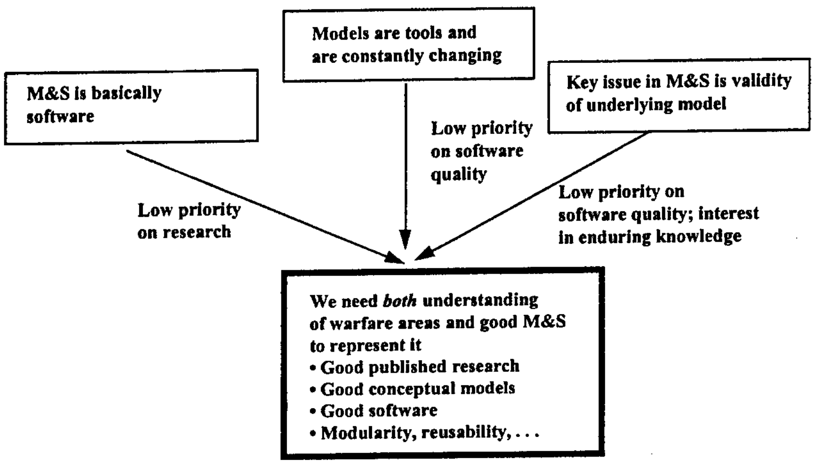

FIGURE 5.1 Synthesis of desires for research base, conceptual models, and effective M&S software.

M&S with data structures that do not relate well to experimental data or to changes of perspective.[3]

Modularity of Knowledge

Another reason for emphasizing research rather than model building per se is that if one attempts to build a comprehensive model of complex systems, there is a good chance of failure: the computer model will eventually collapse under its own weight. By contrast, modular knowledge can endure. Often, specialists are

[3]This issue of databases not being easy to change, is connected to the difference between *declarative and procedural knowledge* (a distinction much discussed in the computer science and artificial intelligence literatures). Declarative knowledge often takes the form of relationships (conservation laws, Newton's laws, and the like), whereas procedural knowledge usually takes the form of a recipe-like method for solving a problem or operating some system (e.g., "After 6 p.m., turn on the lights"). In many respects, declarative knowledge is more powerful, because it can be used to address a wide range of situations. In contrast, procedural knowledge is often "brittle" (at some times of year, sunset may not be until 9 p.m. or so). When we speak of computers lacking common sense and humans being more adaptive, one of the underlying considerations is that computers are typically programmed to be extremely literal, while humans are able to draw on more general considerations to tailor actions to the task at hand.

A quintessential example is the difference between mission orders and detailed instructions. A commander can specify objectives, describe issues and constraints, and then let his subordinates achieve those as proves feasible and appropriate. Alternatively, he can lay out a plan that "scripts" their activities. The former expresses the problem with declarative knowledge; the latter with procedural knowledge.

needed to understand different aspects of system behavior (e.g., probabilistic issues versus the effects of saltwater on instrumentation versus the location errors associated with a less-than-complete GPS constellation).

PRIORITIZING WARFARE SUBJECTS FOR RESEARCH

Calling for an across-the-board program of research would be of little value. Further, the panel is recommending a significant change in the way business is done, which raises the barriers. With this in mind, the panel identifies a first set of warfare subjects for priority attention by the Navy and Marines. Success in these domains might lead to more general changes later.

In developing this priority set, the panel established several criteria:

- Military importance to the Navy or Marine Corps,
- Importance to higher-level joint operations,
- High potential payoff for empirical and theoretical research, and
- Need for interest in and oversight by "operators" and "warriors," rather than scientists alone.

With these criteria as background, Table 5.1 provides a possible first list of subjects. Each has *major* knowledge gaps that could be narrowed by empirical and theoretical research closely tied to the "warrior communities."

DESIRED ATTRIBUTES OF RESEARCH PROGRAMS

Although the research needed would obviously vary from subject area to subject area, the following features would seem to be strongly desirable in most cases. An overarching theme is the need to take a holistic approach rather than one based on either top-down or bottom-up theology. For each warfare area the panel recommends developing hierarchically integrated families of models with different characters and resolutions—not to predict detailed behaviors, but rather to explore and understand military phenomena. Such simulation-based exploration is a form of experimentation that can yield profound insights.

1. *Top-down thinking from the perspective of a JTF commander.* A common failure of warfare research, and of model building, is not representing from the outset some of the principal factors that would affect higher-level operational decisions. As a consequence, even detailed models often have an overly narrow scope. Further, there are forest-and-trees problems. This can be mitigated by defining the problem initially so as to include at least primitive representations of higher-level strategy and command-and-control.

2. *Tactics development.* Whenever one considers something new (e.g., a new operational concept or weapon system), it is important to think about the

TABLE 5.1 Warefare Areas Needing Empirical and Theoretical Research

Warfare Area	Shortfalls in Knowledge of Phenomena	Importance	Potential Value of Focused Research	Comments
Joint task force operations with dispersed forces	•••	•••	•••	Issues of survivability and effectiveness (may need probabilistic depictions)
Effectiveness of long-range precision strike against armies taking countermeasures	••	•••	•••	Likely large differences among weapon enthusiasts, planners, and on-the-ground reality
Short-notice early-entry operations against opposition	•	••	•••	Short-notice planning and mission rehearsal
Theater-missile defense, including counterforce, and including speed-of-light weapon options	••	•••	•••	M&S will be only mechanism for evaluating effectiveness in large-scale battles
Expeditionary warfare and littoral operations	••	•••	•••	Problems with smart mines, opposition, missiles, and WMD

range of possible tactics—for both oneself and the adversary, perhaps through some cycles of measure and countermeasure. Historians repeatedly remind us that the principal changes wrought by previous revolutions in military affairs (RMAs) have been at least as much organizational and doctrinal as technical.

3. *Realistic decision behaviors.* In each domain, there is likely to be a need to build realistic decision models, for both friendly and adversary sides (and third parties as well), as well as models that represent important limiting cases such as optimization (feasible only with perfect information, but an important bound) and doctrinal behaviors. In addition, the panel believes it important—in both research and the subsequent M&S—to allow for human play. Even so-called "constructive models" for analysis should have human play options. Failure to allow for this will often guarantee slipping into a pattern of computer-comfortable but unimaginative and unrealistic behaviors.

4. *Theory development for multi-resolution model families.* This issue is discussed elsewhere in the report (Chapter 6 and Appendix E), but it should be a key element of research in most warfare areas.

5. *Well-designed empirical work exploiting high-resolution simulation, DIS-modulated military exercises, and training experiences.* It is generally agreed that higher-level M&S lacks an adequate empirical foundation, but it is often suggested that little can be done about the matter. This is no longer true. At least two developments have changed what is possible: (a) the advent of high-quality, high-resolution simulation with extraordinary computer power, and (b) the advent of distributed simulation, including distributed interactive simulation (DIS) as a core activity of training and exercising. With respect to the latter, it is now possible to collect operational data unobtrusively. Figure 5.2 sketches this notion (Davis, 1995b).

6. *Stochastic and human-in-loop options at all levels of play.* In many of the important warfare domains, decision makers—including commanders about to send their people into battle—need to understand the stochastic nature of events. Expected value models can be exceedingly misleading. Or, when their shortcomings are obvious, they can inappropriately discredit otherwise valuable simulation (e.g., as when a fractional carrier is still critical in an operation's success). The panel therefore believes it essential that research attend seriously to such matters. One aspect of this will, again, involve including humans in the loop for at least some of the activities.

7. *Comprehensive models, including "soft" effects.* If simulations are to be realistic, whether or not precise, it is essential that they reflect a vast range of soft factors that include, for example, random human errors, virtual attrition (as when pilots achieve poorer air-to-ground performance when flying in an intense air-defense environment, even if they take no attrition), "frictional" effects, and suppression of enemy effectiveness by information warfare or barrage bombing. While these are sometimes notoriously difficult to incorporate precisely, the

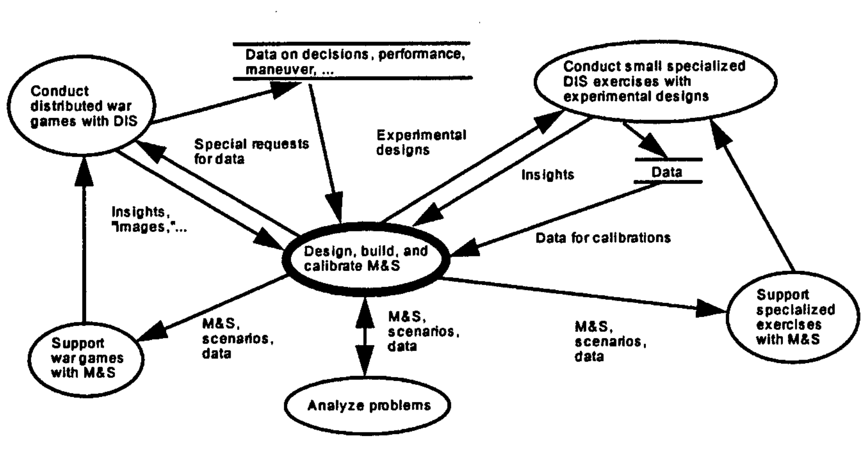

FIGURE 5.2 Exploiting DIS experiments for empirical information. SOURCE: Reprinted, by permission, from Davis (1995b). Copyright 1995 by IEEE.

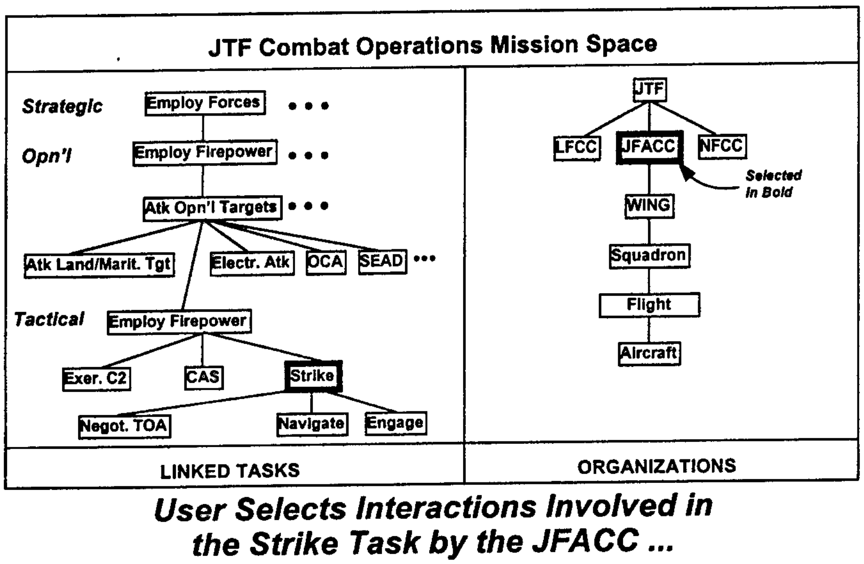

FIGURE 5.3 An illustrative slice of a CMMS for JTF strike operations. SOURCE: Jefferson, DMSO (1996).

principle to be kept in mind is that to omit them is to assume implicitly that they have no effect (i.e., to assume that model correction factors are all unity).

8. *Complex MOEs.* Although operations researchers and mathematicians often like to identify a single measure of effectiveness (MOE) on which to focus, there are relatively few circumstances in which that is appropriate for higher-level decision support. JTF commanders, for example, must worry not only about damage caused to opponent forces, but also about impacts on opponent effectiveness. They must worry even more about effects on the success of their own strategy. And they must typically worry about casualties—to their own forces, to allies, and even to enemy forces. Thus, research and analysis should in the panel's view increasingly provide a rich set of MOEs.

9. *Empirical data development.* Empirical work using DIS exercises is mentioned above, but much broader activities are possible. These include historical research, dispatching operations researchers to observe and report on operations and doctrinal experiments, structured interviewing of military experts—foreign as well as American—and field tests.

10. *Models as products.* In the past, military research has often not connected well with the needs of the M&S community. While it is evident that the panel does not recommend focusing scientists on the building of large-scale M&S, it does believe that more can be demanded of them in terms of expressing their conclusions in the form of discrete models (mathematical, logical, or com-

puterized) that connect conceptually to the larger realm of M&S. Here the panel believes that it is important that military scientists inform and be informed by the emerging work on common models of the mission space (CMMS). As an example here, someone conducting research on command and control might need to see how his ideas could be modularized so as to be useful in a DOD-wide model repository constructed with the imagery of Figure 5.3 in mind.[4]

[4]Adapted from a MORS briefing on CMMS issues sponsored by the Defense Modeling and Simulation Office (1996).

6

Creating and Improving Intellectual and Technological Infrastructure for M&S

KEY TECHNICAL PROBLEMS REQUIRING INVESTMENT

Whereas the "content" of M&S will best be improved with research programs organized around warfare areas, there are some important cross-cutting technical problems that merit separate investment. Some involve modeling theory; some involve infrastructure technology and standardization. Three seem particularly significant in thinking about achieving long-term visions:

- Understanding and M&S of complex systems,
- Families of models, and
- M&S infrastructure.

The first, complex systems, is discussed in Chapter 3. (Appendix B is a much more extensive treatment.) Here the focus is on the second and third.

HIERARCHICALLY INTEGRATED FAMILIES OF MODELS

The first subject involves integrated families of models. (See also Appendix E.) Having such families is important in all domains of M&S. To take merely one example, a JTF commander needs to work for the most part with a highly aggregated view of the theater and forces. However, he also needs to be able to zoom in on particular regions or operations, perhaps because they are critical and must therefore be understood in detail. As a practical matter, all this requires different models (not just a single high-resolution model) because of both complexity and data uncertainty.

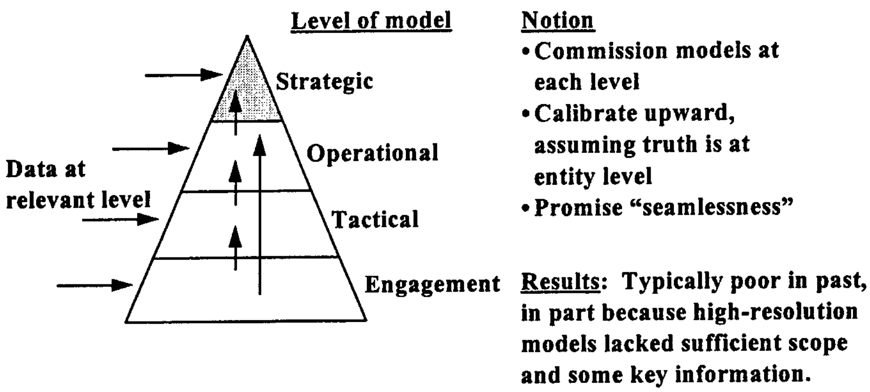

FIGURE 6.1 Old-think on model families.

In one sense, families of models have been around for years, but mostly on viewgraphs. In old-think (Figure 6.1), moreover, they were formed by *legislating* that existing models at different levels of resolution would be considered a family and that detailed models would be used to generate data calibrating less-detailed models.

The results of most efforts along this line have been disappointing, if not downright failures. First, the models declared to be family members often were only casually related. Connecting them proved difficult and ambiguous. In part as a result, but also because of flawed theory, the calibration efforts failed. High-resolution models, for example, often predicted attrition and movement rates that greatly exceeded observed reality—presumably because they were not yet sufficiently complete to reflect many of the delays and other frictional effects that occur in real military operations.[1] Also, the high-resolution models often did not address key features of the problem. That is, they had insufficient scope. In other cases, the high-resolution models were credible, but the low-resolution models had no "hooks" for reflecting the high-resolution results. For example, they depended only on deterministic averages of higher-resolution phenomena when statistical or distributional information was critical. The general problem is that models that have not been designed for cross-calibration are often difficult to relate to one another.

[1] As an example of the difficulties here, suppose that one wants to use a high-resolution simulation of company-level battle to calibrate the attrition rates of a higher-level model. A company, once it is in battle, may have a very short but intense period of attrition. However, most of the time such a company is not in such a battle. Further, there may be many hours of preparation before any such battle occurs. Accounting for these matters in attempting to provide "average rates" remains extremely difficult conceptually and was beyond the simulation and computational states of the art in past decades.

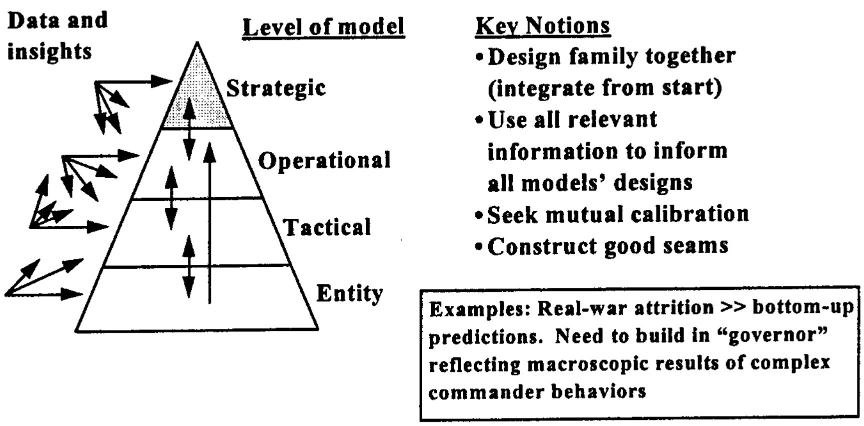

FIGURE 6.2 New-think: integrated hierarchical families of models.

There have also been organizational problems. If the different models of such a hierarchy are owned by different organizations that only occasionally work together, the linkages are more imagined than real, and sometimes cynically constructed when real at all. This is a harsh judgment, and there have been some notable partial successes, but the panel believes the judgment is correct.[2]

Figure 6.2 suggests an image of "new-think" on these matters. Although it may appear "common-sensical," it represents a drastically different image than the one followed in the past and assumed appropriate by most in the analytic community. In this image, models at different levels of detail are designed together from the outset so that there is a true integration. Variables from one level "understand" variables at another. Second, models at any given level are designed to make use of data from other levels of resolution. Returning to the attrition example, if we know from historical evidence (and common sense) that attrition is self-limiting because commanders will not tolerate excessive attrition, then someone building a high-resolution model may need to design in corresponding decision rules that could be calibrated against macroscopic information on behaviors (which might be different for different nations' commanders and forces).

The main point here is that in building models and calibrating them we should be using all the knowledge available, regardless of resolution, and one should be attempting to make the family members consistent with each other.

[2]For a review of such matters, see Davis and Hillestad (1993a,b). The latter mentions two efforts, one by the U.S. Air Force and one by the German IABG, that were reasonably successful in developing model families. Both efforts were tightly managed and were within a single organization. The fundamental difficulties in this domain are now recognized by the Defense Modeling and Simulation Office (DMSO) and DARPA.

One should not assume that "truth" resides at high resolution, low resolution, or indeed at any one level or in any one characterization. For example, current high-resolution, entity-level simulations often contain rich information on microscopic behaviors, but they have very limited scope and no or inadequate representation of higher-level context (e.g., the JTF commander's objectives, strategy, and constraints). By contrast, that information may be readily seen in more aggregate representations of the same war. Connection and calibration, then, should be two-way.[3] This type of thinking is familiar in some types of engineering, but it is quite unusual in combat modeling.

Unfortunately, no one today knows how to carry out the vision of "new think." Doing so will require fundamental research as well as applied research on particular problems. There are existence proofs for such models in relatively simple cases, and the beginnings of a theoretical foundation, but there are many theoretical obstacles.[4]

These matters are discussed more fully in Appendix F. Let it suffice here that operationalizing the ideas suggested in Figure 6.2 is very difficult as a matter of theory. Although there have been many claims to the effect that object-oriented programming creates hierarchies of models, such programming usually focuses on the entities (e.g., corps, division, brigade, battalion, company, platoon, squad). To be sure, such hierarchical entities can be represented more easily in object-oriented programming than older methods, but the more serious representational problems involve *processes* (e.g., attrition, movement, and command-and-control) rather than entities. Relating processes at different levels of aggregation or resolution is conceptually very difficult, and very few military researchers have even attempted to do so rigorously. This is a subject for serious theoretical research.

One Element of Doing Better: More Ambitious High-resolution Simulations

As discussed above, one of the most serious past difficulties in trying to calibrate upward has been that the high-resolution models had insufficient scope and, even within the scope dealt with, incomplete information. For example, in

[3]For a simple example of the two-way calibration issue involving maneuver warfare of ground forces, see Davis (1995a), which works out the problem analytically. For other discussions of multi-resolution modeling issues, see articles by Davis and Hillestad in the edited collection of the Military Operations Research Society (MORS) by Bracken et al. (1995).

[4]See Davis and Hillestad (1993a) for the report of a workshop on variable-resolution modeling sponsored by DARPA and DMSO. One concept discussed in that workshop is the notion of integrated hierarchical variable resolution (IHVR) modeling. The key point here is that if the models at different resolution of key processes such as attrition are related by hierarchies of variables, it is "straightforward" to define procedures for calibration. These must involve summations and integrals with appropriate statistical weighting for the application at hand (see also Appendix E).

past decades it was not possible to have entity-level simulation extending to division and corps in scope. As a result, the high-resolution simulations focused on, for example, company- or battalion-level combat. The result was that the simulated battles occurred to some extent in a vacuum, without representing the lengthy, complicated preparations and maneuvers that typically precede the battles, much less the associated frictional complications. With increased computational power, however (and with the benefit of improved software engineering), it is now feasible to greatly expand scope. Along with doing so will surely come substantially improved ability to "see" and understand the interrelationship of events at different levels of organization.

A second difficulty with most high-resolution simulations has been their failure to incorporate behavioral models representing decision making at the various echelons. Much of the high-resolution work has employed military officers for command-control, which has its own advantages, but which complicates or precludes some of the activities needed for analysis. This limitation is also being overcome, slowly, as improvements are made in so-called semiautomated forces (SAFOR). There have now been a number of model developments, notably in the United States and Germany, that have advanced the state of the art in such matters. In the decades ahead, this agent-based modeling will improve greatly—given adequate support and high enough standards. At present, many workers are pleased when the models represent stereotyped doctrinal tactics at low levels, but, with time, the models will become increasingly adaptive and will probably have "learning capability."[5]

Currently, the emphasis on SAFOR is at low levels (e.g., company level when dealing with land forces). However, decision models are feasible for all echelons, and some have been demonstrated and even used, with various levels of success.[6] The forecast here is cautiously bullish, even though it is likely that selective human play will always be very desirable, not just to calibrate models, but to ensure the range of innovations and "unusual" behaviors.

There are other potential and important improvements that should be sought in high-resolution models. These include better representation of the environment (haze, smoke, snow, sea state, and so on) and better representation of low-

[5]Some aspects of "learning capability" are by no means exotic, however unusual in modeling. Consider a simulation in which the two forces have imperfect information about each other and about some of the "laws of war" (e.g., rates of attrition and movement). As they engage in operations and "observe" simulated events, they can also recalibrate some of their assumptions. Thus, if one side's doctrine calls for fast movements and the other side's assumes slower movements, then both sides should use the simulation's version of "truth" in making decisions after movement rates have been observed.

[6]The EAGLE model, developed initially at Los Alamos National Laboratory and subsequently by TRADOC and MITRE, orginally used script-based methods from the artificial intelligence community to deal with battalion-level decisions. The CONMOD development at Lawrence Livermore Laboratory was never completed, but included extensive design work and some prototype demonstration of option-

level human behavior (not "decisions" so much as human "behavior"). Much work is needed on both, although there have been notable advances.

A Different Perspective: The Need for New Modeling Approaches

While there are many reasons to believe that high-resolution simulations will be greatly improved in the years to come, there are also reasons to doubt that they will ever be able to generate accurate higher-level "truth" without incorporating information and constraints from higher-level (lower-resolution) information and perspectives. What may be feasible in principle (such as working from Schrodinger's equations to engineering detail) is often not feasible in practice. It is striking that approaches emphasizing "agent-based modeling" with adaptive agents following relatively simple principles and rules sometimes have the ability to "generate" remarkably realistic macroscopic behaviors, and that the same workers accomplishing this had previously worked diligently, but failed, in a more exclusive bottom-up-with-detail approach.[7] Interestingly, some of this work is neither high-resolution nor low-resolution in character, but rather something new—e.g., low-level agents with only a few characteristics and behaviors.

One interesting point here is that what some communities refer to as agent-based modeling with emergent behaviors looks to others very much like what others think of as adding adaptive decision models to traditional simulations. Further, it is likely that the agent-based models built on only a few basic principles will not prove robust enough for decision support (unless the models can be validated against extensive empirical data), in which case it seems even more likely that the two approaches will to some extent converge.

Working Toward a Larger Tool Kit: Models Other Than Simulations

One of the peculiar features of the current discussion of M&S is that the vast majority of discussion is about simulations—so much so that it is sometimes forgotten that other powerful forms of modeling exist. Simulation generates a possible behavior over time of the real system being modeled. One sets initial

---

ally automated large-scale high-resolution simulation. The German Armed Forces University has extensive experience with rule-based and utility-maximizing decision models informed by many years of human play. The RAND Strategy Assessment System (RSAS) of the late 1980s included theater-level and even political-level decision models employing a variety of methods that included adaptive scripts (akin to real-world branched war plans). These models were able to recognize failures and opportunities requiring changes of plan (i.e., changes of adaptive plan). The political models took a world view, had extensive situation-assessment capability, and made plausible decisions about escalation, termination, and change of high-level (theater-level or multitheater) strategy.

[7]A good example of this was reported to the panel by Darryl Morgeson and Chris Barrett of Los Alamos National Laboratory, based on their transportation modeling work. There are numerous other examples (see also Appendix B).

conditions, executes the simulation, and watches a rendition of how the systems may behave. Simulations are well suited to certain types of "what-if?" questions because one merely changes the initial conditions and runs the model again. However, simulations are often very complicated—especially entity-level simulations. They become difficult to control and comprehend. Further, they cannot answer many questions of interest to decision makers, such as "Under what conditions would I be able to . . . ?", or "If I must achieve [some level of performance], how many . . . will I need?"[8]

Yet another problem with simulations is that they are in some cases the antithesis of the reductionism that is so often critical in decision support. They are so rich that one can lose the forest for the trees. This is especially troublesome when the aggregate behavior of the system turns out to be much simpler, and much more easily understood, than one would ever imagine from studying simulation inputs or the outcomes of a few runs. Yet it happens frequently in systems that approach some kind of steady state, or in systems in which many of the complex interactions produce a simpler average behavior that can be discussed in simple terms.

The current fascination for simulations and the need to rebalance this with more effort to use other forms of modeling has been discussed by Herbert Simon (1990). Versions of simulation that "go beyond 'what-if?' questions" by logic programming methods to embed knowledge that allows the simulation to find initial conditions sufficient to meet specified end states have been discussed and pursued by RAND's Jeff Rothenberg and colleagues (see Mattock et al., 1995). Also, there may be a revival ahead of defense economics dependent less on simulations than on simpler spreadsheet-level models, cost data, and decision support tools. Within the larger technical community of universities and industry, it is notable that younger workers are increasingly expert in powerful desktop analytical tools, tools such as Mathematica and Macsyma, which accomplish symbolic manipulation as well as perform many other functions. A new tool called Analytica facilitates analytical modeling in which input parameters have associated probability distributions. It also facilitates the hierarchical modeling.[9] While the panel does not discuss such matters much in this report, it believes they merit more attention.

[8]Complexity is sometimes in the eyes of the beholder. Some simulations (e.g., Janus) depend ultimately on a relatively small number of principles and data with a relatively well defined origin. Further, behavior is sometimes rather easy to understand because it so tied to physical processes. However, from an analytical viewpoint the same type of model may seem very complicated because there are so many variables, especially if one does not uncritically accept weapons-effect data, doctrinal estimates of movement, and so on.

[9]Analytica is a product of Lumina Decision Systems, which licenses underlying software from Carnegie Mellon University. Mathematica is a product of Wolfram Research. Macsyma is sold by Macsyma, Inc.

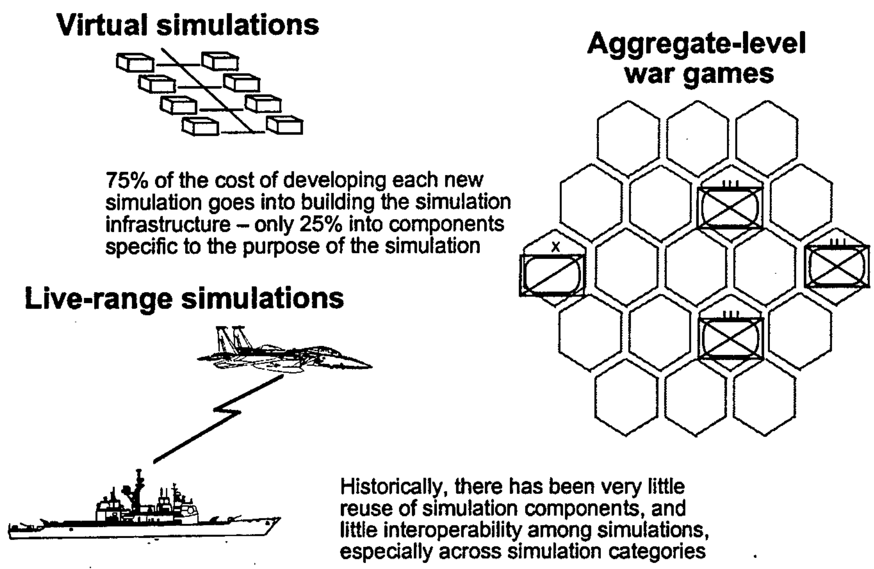

FIGURE 6.3 Rationale for M&S infrastructure.

M&S INFRASTRUCTURE

Rationale

The next subject requiring major technical effort is infrastructure. The panel cannot do justice to this subject here (see also Appendix F).

Figure 6.3 suggests other reasons for supporting infrastructure initiatives. For example, when building a new "stand-alone" simulation, anecdote suggests that it is typical to spend roughly 75 percent of the resources on the underlying infrastructure (e.g., the tedious programming necessary for bookkeeping on entities and for creating interfaces to input/output devices such as databases and graphical displays), and only 25 percent on the specific content that motivated the simulation effort. Although no one should take these figures as precise, there seems to be a consensus on their being roughly right. It should then be no surprise that "new" simulations are often merely a reprogramming of old models, with no substantive improvements. The situation is analogous to expecting improvement in a manuscript by changing the word-processing system. Historically, there has been only relatively little reuse of model components within application classes of M&S (e.g., within the group of constructive models used by a particular organization), and extremely little sharing or reuse across boundaries such as those of Figure 6.3 (virtual simulations, war games, and live-range simulations).

A common infrastructure could yield substantial improvements in development time and productivity. Simulation developers could focus on the specific modules of direct interest to them and to reuse other modules as appropriate.

This approach also facilitates having multiple levels of resolution, levels appropriate to the application. And it minimizes redundancy and inconsistency of simulations developed by different organizations.

Layered Architecture for M&S

To achieve the benefits of a shared simulation infrastructure, a clear architecture is needed. As illustrated in Figure 6.4, this architecture must recognize and address several different layers at which simulations must operate:

- *The computing platform layer*, including the specific workstations or other processors being used to execute the simulations within a federation.
- *The network layer*, which includes the local area networks, wide area networks, and interface modules that permit the computing platforms to communicate efficiently with each other.
- *The simulation layer*, which executes various models to generate the over-

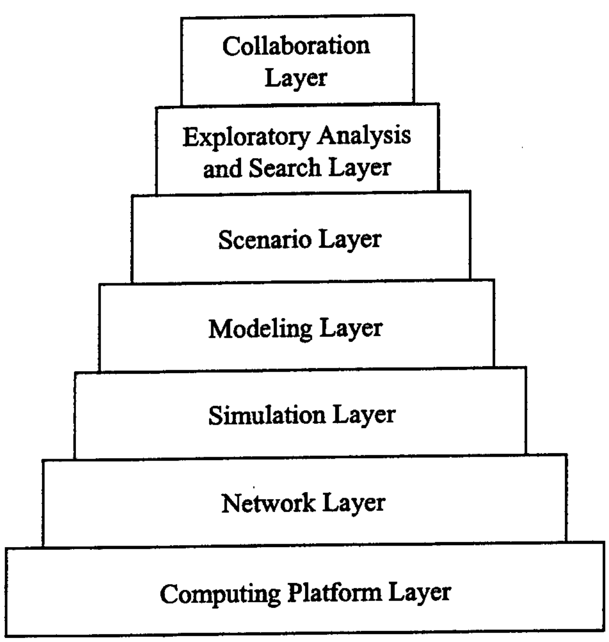

FIGURE 6.4 Layered architecture for M&S.

all simulation behavior that represents the purpose of the study, exercise, rehearsal, or test.

• *The modeling layer*, which includes repositories of models for representing battlefield tactics, weapons, sensors, communications, terrain characteristics, environmental phenomena, and so on. In many cases, new models may need to be developed for a particular application, but the adaptation and/or reuse of previously developed models should often be encouraged.

• *The scenario layer*, which includes the development of force layouts, scripts, and initial conditions relevant for the study, exercise, or rehearsal being planned. Again, it will often be necessary to develop new scenarios for a specific application, but whenever possible, adaptation and/or reuse should be encouraged.

• *The exploratory analysis and search layer*, which supports the exploratory analysis under uncertainty discussed in Appendix B, and also the automated search, where possible, of the system's design space.

• *The collaboration layer*, which electronically supports collaboration among various people involved in an M&S study, enabling them to share data, work on models, and analyze results, and so on, in a coordinated and efficient fashion.

High-level Architecture

The recently promulgated high-level architecture (HLA) for M&S[10] attempts to address some of the issues raised by the layered architecture just presented. HLA is concerned with simulation modularity, interoperability, and component reuse by means of a consistent conceptual approach, domain-independent infrastructure components, and a repository of previously developed simulation modules. All substantive representations of real-world phenomena are maintained inside the simulation components. HLA serves as the "plumbing" that allows the components to interact with each other.

Under the HLA conceptual approach, the set of simulation components that are assembled for the purposes of an analytic study, a training exercise, or a field test is termed a "federation," and the individual components are called "federates." The large majority of federates are simulations, which are responsible for representing some portion of the real-world phenomena under study, but they also include such other components as data collection systems, test status moni-

[10]The memorandum by Undersecretary of Defense Kaminski mandating the high-level architecture, as well as a variety of documents describing it, can be found at the Defense Modeling and Simulation Office's Web site (http://www.dmso.mil). Also, a glossary of M&S terms is available at http://www.dmso.mil/docslib/mspolicy/glossary.html.

toring devices, and controllers' consoles. The latter elements are consumers of simulation data rather than direct participants in the simulation.

One of the fundamental architectural precepts of HLA is that federates interact with each other only through a run-time infrastructure (RTI) in accordance with a well-defined interface specification. The RTI is composed of a number of software modules that provide functional services to the federates. One software module is collocated with each federate. The federates communicate with each other by addressing a service request to their local RTI module and by responding to service requests presented by the local RTI module.

Another precept is that the federation needs to agree on a common object model that includes the types and classes of objects represented, the attributes that represent the state of each object, and the interactions that can be generated by one object to affect the state of another. The HLA defines a format for capturing this information, called the object model template (OMT), and it requires that every simulation that is a candidate for inclusion in a federation must maintain its own simulation object model using the OMT format. In essence, the process of forming a federation consists of a negotiation regarding the various simulation object models, resulting in decisions about which parts of the various simulation object models will be combined to form the overall federation object model. Among the key elements of the federation object model is an agreement about how simulation time will be managed across the federation.

The RTI interface specification defines six groups of services:

1. *Federation management* services provide the basic functions required to control a particular execution of a federation, such as joining and resigning from the execution, starting, pausing, and resuming the flow of time.

2. *Declaration management* services provide the functions by which individual federates convey to the RTI the classes of objects, attributes, and interactions they will represent during a given execution and the classes of objects, attributes, and interactions they need to subscribe to.

3. *Object management* services provide the functions needed to create and delete specific instances of objects of various classes, and to create and delete reflections of the state of objects that are being represented by remote federates.

4. *Ownership management* services provide an opportunity to transfer the responsibility for updating some or all of the attributes of an object to another federate. This permits, for example, the sharing of a high-fidelity sensor output computational capability by several federates.

5. *Time management* services coordinate the advancement of time in a consistent way across the federation. Many modes of time management are supported, ranging from synchronized time steps at agreed-upon rates to negotiations of time advances of arbitrary magnitudes among event-oriented simulations.

6. *Data distribution management* services provide mechanisms for coordinating data publications and subscriptions to ensure the efficient routing of data

only to those federates that have requested it, with minimal amounts of irrelevant data that need to be processed.

The RTI is designed to insulate the individual federates from differences in implementation languages and internal data representations across the federation. Although the HLA makes extensive use of object-oriented representations to describe the interactions among federates, it does not require that any federate use object-orient programming languages or representations internally. Finally, the HLA envisions that federates and their object models will be catalogued in resource repositories where they can be browsed and selected as candidates for reuse in new federations. Resource repository data would include pointers to more detailed documentation and points of contact for those responsible for maintaining various simulation components.

Although HLA makes a substantial contribution to the architecture envisioned in Figure 6.4, it contributes mainly to the lower, technological levels rather than the higher, "intellectual" levels. Indeed, it is somewhat unfortunate that the term "high level" was employed to designate this important development. In particular, while HLA standardizes the simulation infrastructure in which models are executed in distributed fashion, it specifically does not intend to standardize the model content of simulations. However, there are many issues that still need to be addressed in the "modeling layer" of Figure 6.4, and which, if not addressed, could sharply narrow the utility of HLA within DOD. For example, the RTI of HLA usefully contributes to the standardization of time management so that developers need not worry about this aspect of distributed simulation. However, unless there is semantic consistency among models that are being federated, their federation cannot result in a meaningful overall composite.

The Department of the Navy should not only adopt HLA, but also encourage the further development of the higher, "intellectual" layers, where reuse of model content can be facilitated.

Reaction to HLA

General Reactions

As one would expect for any endeavor with such broad potential application and impact, responses to HLA have varied widely. In some circles, it is viewed with considerable suspicion. Much of this suspicion can be attributed to misinformation and natural human tendencies toward fear of the unknown. In some cases, however, one suspects that certain individuals and organizations fear the exposure of the internals of their simulations and possible loss of control over them. In a few cases, there is fear that this approach could lead to the unwise imposition of a few "one-size-fits-all" simulation components that may not be suited for certain applications. And, of course, there is the perennial problem of

a reluctance to pay up-front costs in expectation of benefits that may (or may not) manifest themselves later.

Despite these inevitable concerns and misgivings, HLA is being well received by those with a history of strong commitment to the goal of improved interoperability and reuse. The architecture is inherently broader and more flexible than the distributed interactive simulation (DIS) standards, where many of the advocates of simulation interoperability have historically congregated. The fact that this community voluntarily set aside its DIS standards development activities in order to adopt the HLA is a powerful testimony to the potential of the HLA concept.

As was previously noted, a price must be paid for whatever progress is made. Various M&S user communities must negotiate common conceptual models and definitions that can be used across multiple federations. Decisions must be made about what legacy applications are worth the investment required to overhaul them to bring them into accordance with new standards. Realizing the benefits of improved interoperability and reuse will require the active and unequivocal support of the senior Navy Department leadership.

There will undoubtedly be problems in implementing, promulgating, and institutionalizing these changes. Perhaps some aspects of the current HLA will need to change; perhaps others will be required. The sooner this process gets under way, the better.

As mentioned earlier, in and out of each activity supported by M&S will be flowing not only information, but also models and data. By no means will everything be connected to everything, but substantial reuse and sharing will occur because those doing the work will benefit. And, again, there is more involved here than just model objects and databases. A key element of the M&S infrastructure is commonality of intellectual constructs.

Confusion of Issues: Standards Versus Stamping Out of Variety

Significantly, the panel believes that much of the resistance to HLA is probably due to a confusion of two phenomena. On the one hand, DOD is promulgating content-free standards that should facilitate the marketplace of ideas and products. On the other hand, DOD is constantly exhorting the Services to eliminate alleged redundancies. The image being conveyed is that DOD wants to converge on single models. Indeed, senior officials and military officers have often said as much, although sometimes grudgingly acknowledging that perhaps very modest redundancy (e.g., two models of the same phenomenon?) might be acceptable. That desire to converge on single models is, the panel believes, a serious mistake and quite at odds with the desire to improve the content and general quality of models. The panel will return to this issue later. Here let us merely note that some opposition to the HLA is probably due to the understandable resistance to what is seen as overstandardization of models. By contrast,

HLA developers intend HLA to be content neutral and to facilitate, not obstruct, the marketplace of competitive models.

Let us now make the assumption that the Department of the Navy actively supports HLA and a common infrastructure. Some basic questions will still remain. These include:

- Who develops the simulation modules?
- Who verifies, validates, and accredits these modules for specific uses?
- Who maintains these modules once they have been developed, and updates them as the systems they represent are changed?

Thus, there are many issues ahead.

REPOSITORIES AND MODEL INTEGRATION

It is a waste to have to reinvent the wheel each time a new car is designed. Yet as successive generations of simulations were developed in the past, such wasteful restarts from scratch were the rule rather than the exception. Nowadays, the advent of object-oriented design and programming has provided the technology to support object repositories, where objects may be reused time and time again. These matters are discussed in more detail in Appendix F.

ADVANCED ENVIRONMENTS AND HIGH-LEVEL LANGUAGES FOR M&S

We are all aware of how important "environments" are if we use personal computers. A good current environment allows us to move quickly among applications and transfer material from one application to another—primarily among word processing, graphics, and spreadsheet programs. Also, within each such application we have come to expect tools such as spell checkers, on-line documentation, and hand-holding multimedia primers that lead us through new operations. CADCAM technology is, of course, of great value and becoming well known. So also, then, "environments" are extremely important to both the use and the development of M&S. And high-level languages can be far superior to those pulling users down into levels of programming detail beyond what they need. The commonly used BASIC language with its interactiveness and relatively simple syntax has long been popular for programming by nonexperts. A variety of specialized high-level languages have proved quite powerful to students in science and engineering and to professionals.[11]

[11]Examples here include SIMSCRIPT™ and MODSIM™ simulation languages, the systems dynamics language iThink™, and, much less well known, the RAND-ABEL language™ used to develop the RAND Strategy Assessment System. The programs Mathematica™ and Macsyma™ include high-level language features, as well as facilities for symbolic manipulation and other operations.

Advances in this domain will make it possible to improve greatly the comprehensibility of models, the traceability of results, and the testability in particular contexts. Consider as mere examples here:

- Advanced languages can make it easier for developers to build in "simple explanation facilities," so that an M&S user can see not only the predicted system behavior, but also the key determinants of that behavior. For example, a log statement might say, "Because the JSTARS was inoperative and . . . the acquisition probability for moving targets is reduced by 50 percent from — to — ."
- Further advances in "explanation capability" will make it feasible to query the simulation about why certain events occurred or under what conditions they could occur. Such capabilities would probably depend on logic programming.
- Where terse "explanation log" depictions are inadequate, users should be able to ask for more information and be immediately transported to relevant features of the underlying computer code. If this code is in a high-level language, they may be able to read and understand it directly.
- Or, it may be that what matters are assumptions (i.e., parameter values). Again, with nothing more than a mouse click, the user should be able to see the current values of the relevant parameters—along with documentation about where the data values came from and, in some cases, why they have the value they do (e.g., "based on intelligence reports as of March 28, 2015, it is now believed that the SA-25 surface-to-air missile system as deployed in Libya has the following features : . . ." With another mouse click, the user should be able to read the original intelligence report, which might be posted on-line in DIA headquarters.
- If the user's problem related more to understanding relationships among variables in the underlying model, then he should be able—again with no more than a mouse click and intelligent software noting his context or querying him on the type of information sought—to see a design-level depiction of the model itself. This might take the form of data-flow diagrams, object-model hierarchies, overview text, and so forth, depending on his needs.
- And, of course, it should be possible for the user to "reach back" to model builders, or even to the researchers who provided the knowledge base. This might be done by e-mail, video conferencing phone calls, or a broadcast query to the relevant subset of Web users.

Making complex models comprehensible remains a frontier challenge, and many workers have labored valiantly only to produce simulations understandable only to themselves. Still, there has also been considerable progress. Ironically, there have also been setbacks because of an interesting and frustrating tension between the desires for advanced features and standardization. Currently, work on advanced languages and environments relevant to military modelers seems to have slowed considerably, in large part because those building the advanced tools

need to use methods that are not compatible with commercial software such as Microsoft's Visual Basic™ or the many graphics standards.

This is a passing phase, however, and there will again be major progress. One indication of this is the growing interest in industry-developed methodologies and tools for object-oriented modeling, not just object-oriented programming.[12] Some of these tools are now being used in the JWARS program, for example. They are especially significant because building "explanation capabilities" often depends critically on the clarity and structure of the underlying model design. A related and significant development here is the increasing recognition of the need for common models of the mission space (or CMMS). These can substantially improve the degree to which workers who wish to share each other's models will be able to communicate correctly. That is, they can help improve the semantic interoperability of models.

Finally, the panel notes that commercial industry will probably not support much of what is needed by the Department of the Navy (and DOD) when thinking about comprehensibility, traceability, and the like in combat simulations embedded in command and control systems, or about the competent reuse of models available in a community repository. The incentives for doing so do not yet exist, although we expect that they will emerge in time. Thus, investment is needed. However, its success will probably depend on "squaring the circle," that is, finding ways to incorporate advanced features such as explanation capabilities into software largely written according to emerging industry standards.

RECOMMENDATIONS ON JOINT MODELS

Concerns

One useful focus for Department of the Navy thinking about M&S is the set of joint systems now in development (most prominently JSIMS and JWARS). Taken together, these worthy programs (including the service components) have a price tag approaching $1 billion. It is the DOD's intention that JSIMS and JWARS will become the core for all future joint work on training and analysis, respectively. If successful, JSIMS and JWARS will dominate the joint M&S scene for the next 20 years. Thus, it is important to the Department of the Navy that naval forces be adequately represented. Otherwise, valuable training opportunities will be compromised and the Navy and Marines will suffer in the competitions over doctrinal changes, future missions, and force-structure tradeoffs. More generally, the quality of joint work will suffer.

Unfortunately, it is likely that first-generation versions of JSIMS and JWARS

---

[12]See, for example, Rumbaugh et al. (1991).

will not be satisfactory—even with heroic efforts and even though the products will have many excellent features. There will be major shortcomings with respect to both content and performance. *Consequently, the panel recommends that the Navy insist that DOD and the program offices adopt open-architecture attitudes that will promote rather than discourage substitution of improved modules as ideas arise from the research and operations communities, and that they build explicit and well-exercised mechanisms to assure that such substitutions occur.*

This may seem uncontroversial, and it calls for no more than what some of the programs (notably JSIMS) are projecting, but the history of DOD modeling has often been to produce relatively monolithic and inflexible programs. Further, there has been great DOD emphasis in recent years on avoiding alleged redundancies, collecting "authoritative representations," and exercising configuration control. The panel observed widespread frustration among analysts and other substantive users of models, who see DOD's M&S efforts as driven by civilian and military managers who think models are commodities to be standardized, who sometimes seem to value standardization more highly than quality (harsh words, but too important to be omitted), and who have given near-exclusive emphasis to software technology issues. They and the panel believe M&S should instead be seen as organic, evolving, and flexible systems with no permanent shape (but with standardized infrastructure, including many component pieces).

In fact, the visionary technical infrastructure being promoted by OSD's Defense Modeling and Simulation Office (DMSO) (and software technologists) will permit the open system approach and will permit competition among alternative models (e.g., alternative representations of ballistic-missile defense, mine warfare, or C^4ISR). Thus, while it would be easy for JWARS, JSIMS, and other systems to end up as rigid monoliths, with the right architecture and organizational structure DOD can have its cake and eat it: it can have "standard configurations" while still making it easy for users to substitute model components as new ideas and methods emerge. An important but more subtle aspect of this visionary infrastructure is connecting model evolution to the R&D and operational communities concerned with both current and futuristic doctrine, and, significantly, nurturing a competition of ideas and models. In that way the evolution will be more like survival of the soundest than like continuation of what has previously been approved.

The panel underlines the problem of incorporating research results when they exist because at present the communities who do research and the programming of models often do not communicate well and there is little pressure to assure that the "best" models are reflected in M&S. Indeed, there is much pressure to avoid changes.

Technical Attributes Needed in Joint Models

Against this background of concerns, the panel recommends that the Navy advocate an approach to joint-model development that has a long-haul view and an associated emphasis on flexibility. The groundwork should be in current model-building efforts for the following, which will be important in selected applications in the years ahead:

- *Multi-resolution modeling*, not only of entities, but also of physical and command-control *processes*, with the objective of building integrated models of families with different levels of resolution.
- *Decision modeling to represent commanders at various levels, with both realistic depictions and depictions that provide for optimal behaviors.*
- *Diverse representations of uncertainty*, including use of probability distributions (and, sometimes, alternatives such as fuzzy-set concepts), even in aggregate-level models.
- *Systematic treatment of important correlations* (e.g., the "configural effects" of mine warfare and air defense) (see also Appendix J).
- *Explanation capabilities* linking simulated behavior to situations, parameter values, rules and algorithms, and underlying conceptual models.
- *Mixed modes of play* that are interactive, selectively interruptible (e.g., for only higher-level commander decisions), and automated. (The panel regards the option for human play as critical for analytic applications as well as training, and the option of closed play, for example, of the opponent, as critical for training.)
- *Testing of new doctrinal concepts* requiring new entities, attributes, and processes.
- *Different types of models*. The systems should accommodate model types as diverse as general state-space and simple Lanchester equations, entity-level "physics-based" models, and agent-based models with emergent behaviors. They should employ such varied tools for such uses as statistical analysis, generation of response surfaces, symbolic manipulators, inference engines, and search methods (e.g., genetic algorithms).
- *Tailored assembly. The systems should facilitate tailored creation of models, including relatively simple M&S for specific applications.* That is, one should conceive of JSIMS and JWARS as tool kits with rapid-assembly and modification mechanisms. Excessive complexity is paralyzing and obfuscatory.

In some respects, the last item is the most important. Given the breakthroughs in software technology over the last two decades, it is feasible (though not easy)—and essential—for major M&S efforts to be designed for frequent adaptation, specialization, and module-by-module improvement. One should think of assembling the right model, not taking it from the shelf whole. Further, it should be possible to discard or abstract complexities irrelevant to the problem

at hand. Doing so runs directly counter to the common inclination to seek high resolution for everything, but tailored simplifications are crucial in applications. This is much better understood by those who have used M&S for studies or exercises than by those who develop software. This said, even analysts often find themselves using more cumbersome models than are truly suitable for their purposes. For example, they may use a complex campaign model to examine tradeoffs among deep-strike weapon systems being assessed for their ability to halt advancing armies. Arguably, it would be better to do most of the work with a more specialized and much simpler model with which one could do exploratory analysis.

Finally, a word of caution about the concept of assembly or composition. It is common, in the heady days in which there are more notions and viewgraphs than demonstrated capabilities, for developers to talk loosely of building systems that will be so flexible as to serve quite different functions for distinctly different communities of users. In practice, such visions have seldom proved out. Instead, the systems become so complex—in their effort to serve many user communities—that they do nothing well and working with them becomes difficult and unpleasant. The panel's view is that while designing with an assembly perspective is essential, there are limits to what can be accomplished. Specializations will continue to be needed. It is an open question whether systems like JWARS and JSIMS will prove as versatile as some of the extravagant visions anticipate; DOD's image of their being general-purpose tools may prove wrong. This means that the Department of the Navy (and DOD) should hedge their bets in this regard.[13] One way to do so is to develop stand-alone models for specialized purposes, although perhaps requiring interoperability and the potential for being used within the "big" systems.

RECOMMENDATIONS FOR RESEARCH

Research in Key Warfare Areas

As noted earlier, there has been relatively little recent investment in understanding the phenomenology of military operations at the mission and operational levels. Much of the basis for related M&S is still programmer hypothesis and qualitative opinions expressed by subject matter experts. This has not always been so. During and after World War II, operations research worked from a rich empirical base, but now the United States is entering a period of nonlinear, parallel, information-era warfare for which the intuition of scientists, operations researchers, and warriors is insufficient. Further, it will be relying on complex

[13]The same observation is made within the realm of software. See, for example, Gibbs (1997).

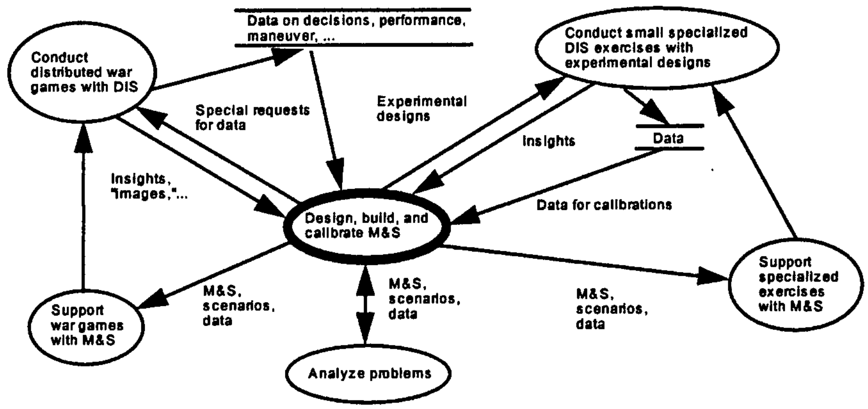

FIGURE 6.5 Using exercises as a source of empirical data for M&S. SOURCE: Reprinted, by permission, from Davis (1995b). Copyright 1995 by IEEE.

systems working as designed in multifaceted joint campaigns. Success may be much less tolerant of errors in concept and execution than in days past.

Subjects of particular importance for M&S-related research in the information era are (1) aspects of command, control, communication, computers, intelligence, surveillance, and reconnaissance (C^4ISR) that involve the content and reliability of information, as well as its transmission; (2) tactics and strategy; (3) human behavior; and (4) the very nature of the extended battlefield in future operations. This list, however, is abstract, and research could easily be disjointed. *The panel recommends that the Navy and Marines select a few high-priority warfare areas and create research programs to support them.* These programs should be organized so as to assure close ties to operational and doctrinal-development communities, and to relevant training and exercise efforts that could be mined as a source of empirical knowledge (e.g., as suggested in Figure 6.5, which would exploit emerging capabilities for distributed interactive simulation).[14] This is a nontrivial and potentially controversial suggestion, since the long-standing tradition has been to avoid—and even prohibit—extensive data collection for use beyond those being trained. The costs of such efforts would be small in comparison with those for buying and operating forces, or even procuring large models. Although the Department of the Navy (and DOD) need to make up for past failures to invest adequately in research, this is a domain in which a total of $20 million to $30 million per year can accomplish a great deal.

As a first list of warfare areas for focused research, the panel recommends the following, which have some overlaps:

[14]Exercises, of course, are another form of simulation—not the "real thing."

- Expeditionary warfare and littoral operations,
- Joint task force operations with dispersed forces,
- Long-range precision strike against forces employing countermeasures,
- Theater-missile defense, including counterforce and speed-of-light weapon options, against very large ballistic-missile and cruise-missile threats, and
- Short-notice early-entry operations with opposition.

Each of the above warfare areas has *major* knowledge gaps that could be narrowed by empirical and theoretical research closely tied to the "warrior communities."

The report describes key attributes of research programs for such warfare areas. An overarching theme is the need to take a holistic approach rather than one based exclusively on either top-down or bottom-up ideas. A second theme is that the research should be seen as focused military science, not model building per se. This will determine the type and range of people involved, and also the depth of the work.

Two examples may be useful here. The first is the challenge of developing command-control concepts for highly dispersed Marine Corps forces operating in small units far from their ship-based support and dependent on a constellation of joint systems. The Marine Corps is studying alternative concepts in the Hunter/Warrior experiments. Such experiments need to be accompanied by systematic research and modeling of different types, perhaps including new types of modeling useful in breaking old mind-sets. It is plausible, for example, that cellular-automata models could help illuminate behaviors of dispersed forces with varying command-control concepts ranging from centralized top-down control to decentralized control based on mission orders. To its great credit, the Marine Corps is currently exploring such possibilities, opting to accept some "hype and smoke" in the realm of controversial complex-system research in exchange for new perspectives and tools useful in doctrinal innovation. While the panel does not believe such simplified models will prove adequate in the long run, they can be very helpful in developing new hypotheses.

A Navy example involves mine and countermine warfare. From prior research based on sophisticated probabilistic modeling accounting for numerous "configural effects" (i.e., effects of temporal and spatial correlations), we know that effective strategies for mine-laying or penetrating minefields are often counterintuitive. By exercising such models and simulation-based alternatives in an exploratory manner (as distinct from answering specific questions), it should be possible to develop decision aids of great value in training, acquisition, and operations. Such aids should not, however, focus only on "best estimate" single-number predictions; they should instead provide commanders with information about odds of success, as a function of information. If the aids are to be useful, they must be informed by an intimate understanding of operational commanders' needs.

7

Challenges in Assimilating and Exploiting M&S Technology

TRADITIONAL CHALLENGES

As noted earlier, M&S is an enabling technology with great potential. However, whenever such a cross-cutting technology is introduced, there are organizational and managerial challenges. It is commonplace for the organization to measure the value of investments against the wrong yardsticks (e.g., saving money in narrow domains, as distinct from changing the very way the organization does business and improving effectiveness for mainstream missions). It is also common for investments to go awry because the new technology is procured and used as an add-on without sufficient buy-in and influence by the organization's core work and workers, because too much is done by committee without leaders and champions who understand the core business, or because the educational groundwork has not been laid.

IMPLICATIONS FOR THE NAVY

Observations

Despite much ongoing success, all of these problems are visible with respect to M&S in the DOD's components, including the Navy and Marines. As examples,

- While substantively broad and ambitious, the Navy's plan for M&S was, as of late 1996, rather "defensive," with viewgraph emphasis on cost savings and

the absence of redundancies. This reflected the attitudes of senior audiences to whom the material was being presented, audiences who look skeptically at M&S.

- There were numerous expressions of concern by office and program directors to the effect that proper investments in M&S would require larger sums of money than available within their own domain alone—even though such investments would have large long-term benefits. This was of particular concern with respect to simulation-based acquisition.
- The Navy's coordination office for M&S has been organizationally weak and may not be well located—especially if the intention is to link M&S to warfighters and decision makers concerned with force structure.
- The perception exists that the Army and, to some extent, the Air Force have "stolen a march" on the Navy in exploiting distributed interactive simulation (DIS) and in laying the groundwork for the revolutionary changes it will make possible. Some of this perception appears to be due to the Army and Air Force having put together coordinating offices and programs earlier and having communicated their efforts broadly. The Navy clearly possesses expertise in DIS and has begun to use it (e.g., the Kernel Blitz exercise and various activities within the Naval Research and Development Division (NRaD), San Diego, California). However, the perception does generally seem to be correct.
- Another aspect of the situation that matters is background. Until recently, at least, the Department of the Navy has been relatively aloof from the last decade's activities with distributed war gaming and advanced distributed simulation. Initially, this stance apparently reflected decisions not to invest in what seemed to be unfocused "technologists-going-crazy" activities. Panel members could understand and to some extent sympathize with that judgment. But the situation is far different in 1997 from what it was a decade ago. In future decades—and surely by 2035—M&S, including advanced distributed simulation, will be altogether ubiquitous and crucial.

The Department of the Navy as a User and Consumer of M&S

A different set of problems relates to the Department of the Navy as a user and consumer of M&S and model-based analysis. The Navy needs to review itself with respect to these matters. While a review of such matters was outside the panel's charge, and the panel drew no conclusions, it notes that there are some troubling reports (see, for example, Calvin et al. (1995)).

If the Navy does have problems being a good consumer of M&S, especially high-level M&S at the mission and campaign levels, the problems have nothing to do with technical training or experience in a broad sense. Indeed, the Department of the Navy has many officers educated and trained in technical specialties such as propulsion and aerodynamics. Further, the Department of the Navy is a generally good consumer of M&S and model-based analysis at the component

level (e.g., physics- and engineering-level work related to weapon system design). The problem, if it exists, seems to be that not enough officers have been focused on the special problems of higher-level M&S, such as those associated with force-on-force analysis, joint-task-force-level analysis, systems analysis, and force planning. It may be that this is in part a consequence of decisions taken in the early 1980s to eliminate OP-96 within the Department of the Navy's organization and to otherwise downgrade the role of operational analysis and its underpinnings, as evidenced, for example, by the then-controversial actions taken regarding the Center for Naval Analyses and its management.

A BACKGROUND OF LEADERSHIP ALOOFNESS FROM M&S

An important background observation is that the Navy's leadership has generally been relatively uninterested in M&S in the past. There were good reasons for this aloofness. Why *should* the leadership pay special attention to M&S as long as components were getting the job done—building ships and aircraft, training, and operating—using models as appropriate? After all, M&S is just a tool. A further consideration here was probably the belief that the M&S technology programs being championed by DARPA were not yet of immediate and great value to the Navy, but could be a sink for money. Yet another concern, which many of the panelists share as the result of their experiences, is that putting too much emphasis on a support activity such as M&S often proves less effective than having core activities "pick up" the support as needed.

Finally, the panel notes that much of this report's discussion focuses on mission and campaign-level (theater-level) models, which historically have been of relatively little interest to senior naval officers because they have generally dealt poorly with naval forces. Deterministic models referring to half a carrier battlegroup launching strike missions seemed inappropriate. Further, there was little treatment in the joint models of the intricacies of antisubmarine warfare, over-the-horizon reconnaissance and surveillance, electronic warfare, the differences in operating conditions in different waters, and so on. The models, then, did not seem very good, but they also posed little danger since experienced military officers and analysts could watch over their use and make corrections for foolishness.

And, to put the matter bluntly, the Navy was not at the center of attention in related studies and exercises dealing with major conflicts: its role was greater in presence operations, peacetime missions of all kinds, strategic nuclear and counter-SSBN operations, and lesser conflicts involving naval forces.

In summary, there have been many reasons for the relative lack of interest in M&S by Navy leadership. Further, the strategy taken may well have been optimal for many years: a strategy emphasizing decentralization and viewing M&S as merely a tool to be used as needed by those responsible for "real" tasks.

NEW CIRCUMSTANCES AND THE NEED FOR TECHNOLOGY-DRIVEN ATTITUDE CHANGES

Against this background, the panel believes that it is time for a fundamental change in the way the Department of the Navy (and DOD) thinks about M&S. Its potential value is enormous for acquisition, training, and operations. However, major problems are arising, and many forecasters are being much too optimistic about how quickly difficulties will be overcome. The issues are more substantive than technological; they involve model content and quality. In the past, models were best seen as mere tools to be developed and used in specific contexts. The people working on problems had also built or overseen the building of the models, were aware of their limitations, and could deal with them. Also, the results could be observed and assessed by military officers drawing on doctrine, lore, and personal experience for intuition about what was and was not credible simulation behavior. That era is passing.

The nature of warfare, and even of peacetime military operations, is changing drastically. Intuition based on traditional doctrine—often dating back to World War II—will no longer be dependable. So it is that all the Services are busily engaged in battle-lab experiments (e.g., the Marines' Sea Dragon activities). One thing is clear, and that is the increased importance of jointness. This is no longer something for lip service, because U.S. military doctrine is now shifting rather dramatically toward a systems-of-systems approach in which jointness is essential (Shalikashvili, 1996). All of this requires extensive simulation, which will become thoroughly embedded in worldwide and theaterwide command-and-control systems. Naval forces have a prime role in many of the envisioned activities, including the theater-opening campaign and theater missile defense.[1] Unfortunately, M&S is becoming extremely complex. The originators of the models will not be present to watch over their use, and many applications of M&S will be made by people who did not develop the models in the first place. For all of these reasons, it is increasingly critical that the most rigorous efforts are made to improve the quality of M&S as judged not only by "validity" (which can be assessed only in context and with discussion of how uncertainty is handled), but also by transparency, flexibility, and in-context testability. At the same time, if the Department of the Navy is to benefit fully from the great potential of M&S—especially in exploiting what is sometimes called the revolution in business affairs—it must arrange for suitable investments.

As suggested in Figure 7.1, taken together, these considerations imply the need for a technology-driven attitude shift represented by a strategic commitment to exploiting M&S technology. This should lead to strategy, policy, and investments. Key elements of such a strategy would include officer education, vigor-

[1]See Naval Studies Board (1996) for discussion of naval forces in regional conflicts. See also the Defense Science Board (1996c).

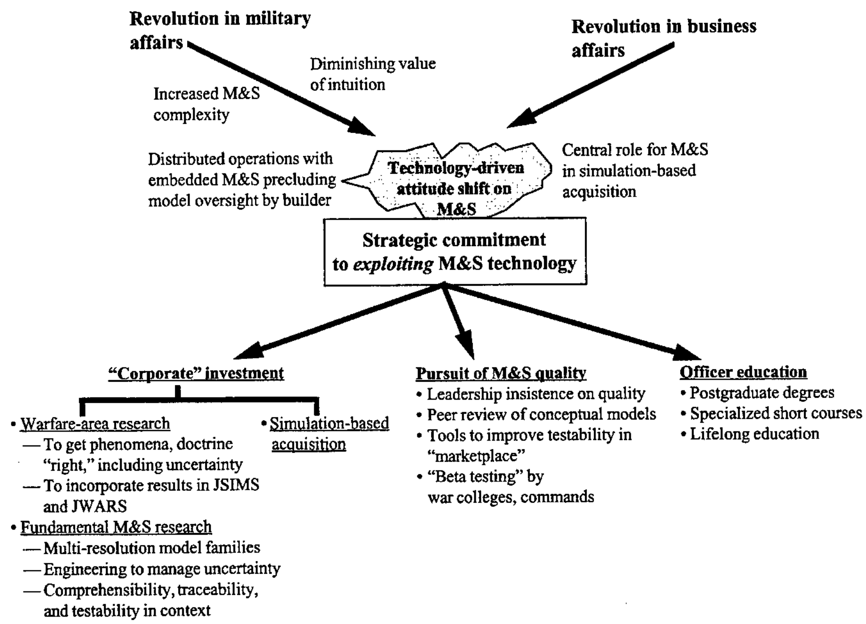

FIGURE 7.1 Despite hype, it is time to look at M&S differently.

ous and rigorous pursuit of M&S quality, research to better understand the phenomena of and doctrine needed for future conflict, and fundamental research on M&S to generate systems that are more understandable and testable. The other element is large-scale investment in simulation-based acquisition, which can—over the next two to three decades—revolutionize acquisition.

ISSUES FOR THE DEPARTMENT OF THE NAVY

Conclusions on Managerial Issues

To reiterate, and go into more detail, the panel concludes that the Department of the Navy needs to make a strategic commitment to M&S. This is not merely a technology study's expression of the view that the Department of the Navy should "pay more attention to M&S." Instead, it is a considered judgment about strategy in the information era. Although the panel members have expertise in M&S, the panel also includes a good deal of expertise in strategic planning and organizational behavior.

What would such a commitment mean? The panel's answer, based significantly on panelists' experience in their own organizations and knowledge of developments in industry, is that such a strategic commitment will mean earmarking capital and continuing resources for the following:

- Simulation-based design, acquisition (major investments),
- Research on the scientific knowledge base,
- Research on M&S and analytic methods, including "simulation science" for complex systems and tools to improve transparency, flexibility, and adaptability of models,
- Education, and
- Links to related joint worlds.

It will also require creating processes for continuing M&S management:

- Appropriate organizational structure,
- Common infrastructure development and support,
- M&S module development and support,
- Ongoing configuration management, and
- Coherent and realistic verification, validation, and accreditation of models and simulations.

More Specific Recommendations

Moving to concrete recommendations, the starting point is to suggest that the Department of the Navy *organize* for a strategic approach to M&S. Box 7.1 itemizes the recommended actions. The panel does not suggest details about organization, except to observe that whoever is responsible for M&S leadership needs to be in a position of influence. The panel also believes that the Chief of Naval Operations (CNO) would be well served by having a dedicated analytical

BOX 7.1
Recommended Managerial Actions

- Organize for effective shepherding through transition.
- Establish crack "analytic team" to serve CNO.
  —Conduct studies.
  —Focus, interpret, and filter M&S.
- Fund key efforts top-down, not by trading within budget categories: focus on innovation and reengineering, not auditing or stamping out redundancy.
- *Support and exploit DOD's emerging M&S infrastructure: Embrace HLA and activities capitalizing on it.
- *Invest in developing and nurturing the knowledge base: Support research tagged for key warfare areas and key technical problems. Couple research to operators.
- Invest in officer education for building, managing, and using M&S (and analysis).

organization. Such an office could also provide intellectual guidance to M&S activities, which too often appear to have a technology-push momentum with only vague notions about how to accomplish objectives supporting core missions.

The panel urges that the Department of the Navy establish funding mechanisms and funding procedures that assure an appropriately "strategic" approach to major issues such as simulation-based acquisition and other activities that would help enable reengineering. This would mean providing some capital investment from the top down, both for long-lead-time investments and common-good investments.

The items marked with an asterisk in Box 7.1 indicate subjects on which the panel focused. For reasons elaborated elsewhere in the report, the panel recommends that the Department of the Navy leadership embrace the infrastructure initiatives associated with DOD's high-level architecture (HLA) and that it specifically endorse and assure research activities to improve the *validity* of models, especially those dealing with combat, logistics, and so on (as distinct from, say, sensor-level models where there is a stronger current knowledge base). The Department of the Navy—and DOD more generally—are severely underinvested here, especially since new-era warfare involves new, complex, and unfamiliar operations. *Major* errors are possible, and scientists, not just computer scientists, are needed here.

A Market Approach to Improving Model Quality and Credibility

Although the panel has discussed verification, validation, and accreditation (Chapter 4), long experience suggests that bureaucratic processes to "enforce VV&A," while potentially valuable, are unlikely to do the job (and could reduce efficiencies). A more effective approach is to exploit natural market forces, that is, by having model users play a more vigorous role. This may sound like nothing new, but the reality is that model users have traditionally found it very difficult to review and test any but simple M&S, except by observing behaviors in some standard cases. Much more can be done with:

- Leadership demanding and valuing continued rigorous attention to M&S quality from the outset of projects;
- Enforced standards for documenting conceptual model (as distinct from the program), the structural relationship between that and the program, and how to review databases (consumers can do little without such basic information);
- Routine use of outside scientific panels to review conceptual models—in part on their own terms and in part to advise the Navy on whether the models and modeling approaches being used exploit the then-current state of expertise;
- Model-building technology greatly improving
  —Transparency and the ability of users to query the system with particular questions,

—The ability to do "exploratory analyses" as a routine part of model testing for a given context, and

—Increasingly competent software agents for verification testing and even domain-specific knowledge testing;

- Routine mechanisms for "beta testing" by appropriate communities, such as the Naval War College and operating commands.

Establishing Research Programs

This report does not recommend any particular organizational approach for assuring the needed research. However, some observations are appropriate. It is important to distinguish between the warfare-area research and the more fundamental research. The latter could rather naturally fall into the domain of the Office of Naval Research and possible Navy-DARPA cooperative efforts. The warfare-area research, however, creates some challenges. Several approaches are possible:

- In each of a set of key warfare areas, require "program managers" (or other relevant managers) to create a research component.
- Create a higher-level cross-cutting office that would support a portfolio of applied warfare-area research in close cooperation with the above managers as well as warfighters and doctrine developers.
- Ask some existing organization or combination of organizations to play this role (e.g., the Center for Naval Analyses and the Applied Physics Laboratory, Johns Hopkins University).

The first approach might maximize the closeness of relationship between researchers and warriors. On the other hand, the fit might not be natural and the research component might get short shrift. Also, the second approach would be likely to have integrative advantages and economies of scale. The third approach may be considered a variant of the second.

Centralization Issues

There are many issues that the panel has not attempted to resolve. One of the more important involves degree of centralization. How centralized should the Department of the Navy management of M&S be? Centralization has advantages. For example, it can improve ability to do comprehensive planning and coordination, but it may also lead to bureaucratic frictions and stifle both creativity and problem solving. Decentralization has great advantages. It allows program managers to address their own problems with minimal external considerations. However, this may inhibit realizing efficiencies and reduce effective

coordination. Centralized funding makes common-good efforts easier, but the resulting programs may be hobby horses, failures, or irrelevant because of being too distant from where the real action is, within the programs and in operations. It seems clear to the panel that the Department of the Navy needs a good deal more centralization than it has today (almost none), but finding the right balance is inherently difficult.[2] It is even more difficult for outsiders to judge.

Education of Officers

A key element of any assimilation effort must be an increased emphasis on educating young officers. The effective exploitation of M&S depends upon the experience, knowledge, and wisdom of its practitioners, hence upon their education. The panel recommends Navy investment in such education at all levels: for those who acquire and design M&S tools and also for those who rely on them to guide acquisition, training, and operations. Some of the education should be in the form of enhanced master's and Ph.D.-level programs. Other aspects should include short courses tailored for officers needing refresher courses, technology updates, and preparation for next assignments involving M&S management. One new educational activity is a master's degree program at the Naval Postgraduate School with OPNAV endorsement. It will emphasize both computer technology (e.g., virtual simulation) and also human-computer interaction modeling, with a strong component of operations analysis. This and other programs—and a competition of programs is important—could make a significant difference over time.

Significantly here, the panel does not necessarily recommend an emphasis on computer science alone, but rather a priority on increasing the supply of young officers with rigorous training in the "hard" sciences or engineering that includes solid exposure to modern M&S, including software engineering. It is common for crack teams in industry doing projects with advanced M&S to be composed mostly of engineers, mathematicians, operations researchers, and scientists.

Realism suggests that master's-level education is much more likely to create wise consumers than practitioners. Thus, the Navy will wish to consider how many of its personnel should go on to obtain Ph.D.s and how best to link up with the best expertise in the university and private domains.

[2]The panel was divided on the recommendations of Calvin et al. (1995) regarding model management. Panel members were sympathetic to this Center for Naval Analyses report's diagnosis of problems (e.g., chronic inattention to verification and validation of models, and inadequate analytical sophistication regarding what model-base analyses are and are not sound), but there was considerable doubt about the prescriptions, which seemed to some of the panel to call for an excessively centralized approach that would generate bureaucracy and associated frictions.

Education of Analysts and Model Builders

Interestingly, the challenge is not just to educate more naval forces officers in M&S, but also to educate more of the analysts and model builders in military operations. This has seldom been recognized as a requirement, but it seems quite important to the panel. One mechanism for doing this is to strongly support the DMSO's efforts on common models of the mission space (CMMS), something that the JSIMS and JWARS programs are currently cooperating on. Other mechanisms could be more formal. For example, the panel can imagine the Naval War College taking a lead role in arranging orientation meetings in doctrine, field visits, and relatively in-depth discussion of the emerging CMMS.

Bibliography

Allen, Patrick. 1992. *Situational Force Scoring: Accounting for Combined Arms Effects in Aggregate Combat Models*, N-3423-NA, RAND, Santa Monica, Calif.

Allen, Patrick. 1995. "The Need to Represent a Wide Variety of Battle Types in Air-Ground Combat Models," *Military Operations Research,* 1(3):19-25.

Almasi, George S. and Allan Gottlieb. 1989. *Highly Parallel Computing,* Benjamin-Cummings, Redwood City, Calif.

Ancker, C.J. 1995. "A Proposed Foundation for a Theory of Combat," *Naval Research Logistics,* 42(3):311-344, April.

Anderson, Robert and Paul K. Davis. 1993. *Proceedings of a Workshop on Methods and Tools for Verification, Validation, and Accreditation*, PM-179-DMSO, RAND, Santa Monica, Calif., October 5-6.

Appleget, Jeffrey A. 1995. "The Combat Simulation of Desert Storm with Applications for Contingency Operations," *Naval Research Logistics*, 42(4):691-714, June. Also in Bracken et al. (1995), pp. 549-572.

Bankes, Steven C. 1993. "Exploratory Modeling for Policy Analysis," *Operations Research*, 41(3), May-June, pp. 435-449

Bankes, Steven C. 1996. "Computational Experiments for Decision Making in Complex and Uncertain Environments," *Proceedings of the Sixth Annual Conference on AI, Simulation and Planning in High Autonomy Systems*, The University of Arizona, Tucson, Ariz.

Bennett, Bruce W. 1995. *Observations on Verification, Validation, and Accreditation (VV&A)*, PM-462-OSD, RAND, Santa Monica, Calif.

Bennett, Bruce, Daniel Fox, Carl Jones, Larry Painter, John Schrader, Robert Weissler, and Barry Wilson. 1992. *RSAS 4.6 Summary*, N-3534-NA, RAND, Santa Monica, Calif. (There is unpublished documentation of the 1993 version of the RSAS, RSAS 5.2.)

Bennett, Bruce, Arthur Bullock, Daniel Fox, Carl Jones, John Schrader, Robert Weissler, and Barry Wilson. 1994. *JICM 1.0 Summary*, MR-383-NA, RAND, Santa Monica, Calif.

Biddle, Steven. 1996. "Victory Misunderstood: What the Gulf War Taught Us About the Future of Conflict," *International Security*, 21(2):139-179, Fall.

Bitters, David L. 1995. "Efficient Concentration of Forces, or How to Fight Outnumbered and Win," *Naval Research Logistics*, 42(3):397-418, April. Also in Bracken et al. (1995).

Bonder, Seth, and L.W. Farrell (eds.). 1970. *Development of Models for Defense Systems Planning*, SRL 2147 TR70-2, Systems Research Laboratory, Ann Arbor, Mich.

Bonder, S., M. Farrell, and K. Close. 1994. *An Analysis of Doctrinal and Tactical Alternatives for Information Campaigns*, three volumes, VRI-G-94-108, Vector Research, Inc., Ann Arbor, Mich., June.

Bos, George E.P., William G. Hunter, and J. Stuart Hunter. 1978. *Statistics for Experimenters*, John Wiley & Sons, New York.

Bracken, Jerome. 1995. "Lanchester Models of the Ardennes Campaign," *Naval Research Logistics*, 42(4):559-578, June. Also in Bracken et al. (1995).

Bracken, Jerome, Moshe Kress, and Richard E. Rosenthal (eds.). 1995. *Warfare Modeling*, published for the Military Operations Research Society, John Wiley and Sons, Danvers, Mass.

Calvin, Edward, Francoise Becker, John Bentrup, and Scott Bailey. 1995. *Developing a Navy Strategy for Modeling and Simulation: Final Report*, CRM 94-196, Center for Naval Analyses, Alexandria, Va., March.

Chew, Jennifer. 1997. "Testing and Validation of Models and Simulations," *Phalanx*, 30(2):8, June.

Cho, T.H., and B.P. Zeigler. 1997. "Simulation of Intelligent Hierarchical Flexible Manufacturing," *IEEE Trans. System, Man and Cybernetics*, 27(1):116-125.

Committee on National Statistics, National Research Council. 1995. *Statistical Methods for Testing and Evaluation of Defense Systems, Interim Report*. National Academy Press, Washington, D.C.

Courtois, P.J. 1985. "On Time and Space Decomposition of Complex Structures," *Communications of the ACM*, 28(6):590-603.

Danneberg, Klaus and Patricia Sanders. 1997. "New Directions in Modeling and Simulation," *Aerospace America*, 35(1):34-39, January.

Davis, Paul K. 1992. *Generalizing Concepts and Methods of Verification, Validation, and Accreditation (VV&A) for Military Simulations*, R-4249-ACQ, RAND, Santa Monica, Calif.

Davis, Paul K. 1993. *An Introduction to Variable-Resolution Modeling and Cross-Resolution Model Connection*, R-4252-DARPA, RAND, Santa Monica, Calif. Also appears in the *Naval Logistics Journal*, 42(2):151-182, March (1995) and in Bracken et al. (1995), pp. 5-36.

Davis, Paul K. (ed.). 1994. *New Challenges in Defense Analysis: Rethinking How Much Is Enough*, RAND, Santa Monica, Calif.

Davis, Paul K. 1995a. *Aggregation, Disaggregation, and the 3:1 Rule in Ground Combat*, RAND, Santa Monica, Calif. Web site at http://www.rand.org.

Davis, Paul K. 1995b. "Distributed Interactive Simulation in the Evolution of DOD Warfare Modeling and Simulation,"*Proceedings of the IEEE*, 83(8):1138-1155, August.

Davis, Paul K., and James A. Winnefeld. 1983. *The RAND Strategy Assessment Center*, R-2945-DNA, RAND, Santa Monica, Calif.

Davis, Paul K., and Donald Blumenthal. 1991. *The Base of Sand: A White Paper on the State of Military Combat Modeling*, RAND, Santa Monica, Calif.

Davis, Paul K., and Richard Hillestad. 1992. "Using Simulation in the Education of General Officers," *Proceedings of the 1992 Summer Computer Simulation Conference*, Society for Computer Simulation, San Diego, Calif., pp. 1071-1076.

Davis, Paul K., and Richard Hillestad. 1993a. "Families of Models That Cross Levels of Resolution: Issues for Design, Calibration, and Management," *Proceedings of the 1993 Winter Computer Simulation Conference*, Association for Computing Machinery, New York, December, pp. 1003-1012.

Davis, Paul K., and Richard Hillestad (eds.). 1993b. *Proceedings of Conference on Variable Resolution Modeling*, CF-103-DARPA, RAND, Santa Monica, Calif.

Davis, Paul K., and Robert Howe. 1990. *Representing Operational Strategies and Command-Control in the RAND Strategy Assessment System (RSAS)*, N-3232, RAND, Santa Monica, Calif.

Davis, Paul K., and Reiner K. Huber. 1992. *Variable Resolution Modeling: Issues, Principles and Challenges*, N-3400-DARPA, RAND, Santa Monica, Calif.

Davis, Paul K., David Gompert, and Richard Kugler. 1996. *Adaptiveness in National Defense: The Basis of a New Framework*, Issue Paper IP-155, RAND, Santa Monica, Calif. A variant of this can be found, along with additional analytical material, as chapters 4 and 6 of Khalilzad and Ochmanek (1997), an edited collection of RAND work done for the Quadrennial Defense Review (QDR).

Defense Modeling and Simulation Office. 1995a. *Department of Defense Modeling and Simulation (M&S) Master Plan*, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C., October.

Defense Modeling and Simulation Office. 1995b. "Glossary of Modeling and Simulation (M&S) Terms," draft, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C., October. Web site: http://www.dmso.mil/docslib/mspolicy/glossary/glossary.html.

Defense Modeling and Simulation Office. 1996a. *Department of Defense Verification, Validation, and Accreditation (VV&A): Recommended Practices Guide,* Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C., November.

Defense Modeling and Simulation Office. 1996b. "DOD High Level Architecture for Modeling and Simulation," briefing by Judith Dahmann, December. Available at Web site: www.dmso.mil, and updated from time to time.

Defense Modeling and Simulation Office. 1996c. *High Level Architecture Interface Specification*, Version 1.0, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C.

Defense Modeling and Simulation Office. 1996d. *Modeling and Simulation (M&S) Master Plan*, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C. Available at Web site: http://www.dmso.mil/docslib/mspolicy/msmp/1095msmp/.

Defense Science Board. 1995. *Defense Science Board 1995 Summer Study Task Force on Investments for 21st Century Military Superiority*, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C.

Defense Science Board. 1996a. "Appendix B: Simulation Technology for the 21st Century," *Task Force on Tactics and Technology for the 21st Century Military Superiority, Volume 2, Part I: Supporting Materials*, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C., October, pp. 75-83.

Defense Science Board. 1996b. *Understanding Distributed-Force Concepts for Rapid Deployment Options,* Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C., November.

Defense Science Board. 1996c. *Tactics and Technology for 21st-Century Military Superiority*, Department of Defense, Washington, D.C.

Deitchman, Seymour. 1962. "A Lanchester Model of Guerrilla Warfare," *Operations Research,* 10:818-827.

Dewar, James, Steven Bankes, James Hodges, Thomas Lucas, D.K. Saunders-Newton, and Patrick Vye. 1996. *Credible Uses of the Distributed Interactive Simulation (DIS) System*, MR-607-A, RAND, Santa Monica, Calif.

Director, Defense Research and Engineering. 1996a. *The High Level Architecture for Modeling and Simulation*, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C.

Director, Defense Research and Engineering. 1996b. *Joint Warfighting Science and Technology Plan*, Office of the Under Secretary of Defense for Acquisition and Technology and the Joint Chiefs of Staff, Washington, D.C.

Dupuy, Trevor. 1987. *Understanding War*, Paragon House Publishers, New York.

Epstein, Joshua M. 1988. "Dynamic Analysis and the Conventional Balance in Europe," *International Security*, 12(4):154-165, Spring.

Farrell, Robert L. 1989. *Remarks on Ground Combat Attrition Modeling*, paper presented at Army Research Office Workshop on Attrition Modeling in Large-Scale Simulations, February 2-4, Washington, D.C.

Fogel, D.B. 1994. "An Introduction to Simulated Evolutionary Optimization," *IEEE Transactions on Neural Nets*, 5(1):3-14, January.

Fujimoto, R. 1990. "Parallel Discrete Event Simulation." *Communications of the ACM*, 33(10):30-53.

Gard, Thomas C. 1988. "Aggregation in Stochastic Ecosystem Models," *Ecological Modeling*, 44:153-164.

Gibbs, W. Wayt. 1997. "Taking Computers to Task," *Scientific American*, 277(1):82-89, July.

Goldberg, D.E. 1992. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, Mass.

Gorges-Schleuter, Martina. 1989. "ASPARAGOS: An Asynchronous Parallel Genetic Optimization Strategy." *Proceedings of the 3rd International Conference on Genetic Algorithms*, J. David Schaffer (ed.), Morgan Kaufmann Publishers, Inc., San Mateo, Calif., pp. 422-427.

Gye, Richard, and Toby Lewis. 1986. "Lanchester's Equations: Mathematics and the Art of War," *Mathematical Scientist*, 1:107-119.

Hartley, Dean S., III. 1995. "A Mathematical Model of Attrition Data," *Naval Research Logistics*, 42(4):585-608, June.

Hartley, Dean S., III, and R.L. Helmbold. 1995. "Validating Lanchester's Square Law and Other Attrition Models," *Naval Research Logistics*, 42(4):609-634, June.

Helmbold, Robert. 1983. "Combat Analysis," *Phalanx*, 16(11-15), March.

Helmbold, Robert, and Alan S. Rehm (translators). 1995. "The Influence of the Numerical Strength of Engaged Forces in Their Casualties," by M. Osipov (1915), *Naval Research Logistics*, 42(3):435-490, April.

Hillestad, Richard, and Mario Juncosa. 1993. *Cutting Some Trees to See the Forest: On Aggregation and Disaggregation in Combat Models*, R-4250, RAND, Santa Monica, Calif.

Hillestad, Richard, and Louis Moore. 1996. *The Theater-Level Campaign Model: A Research Prototype for a New Type of Combat Model*, MR-388, RAND, Santa Monica, Calif.

Hillestad, Richard, John Owen, and Donald Blumenthal. 1992. "Experiments in Variable Resolution Combat Modeling," N-3631-DARPA, RAND, Santa Monica, Calif. Also published in *Naval Research Logistics*, 42(2):209-232, March (1995), and in Bracken et al. (1995), pp. 63-86.

Hillestad, Richard, Bart Bennett, and Louis Moore. 1996. *Modeling for Campaign Analysis: Lessons for the Next Generation of Models, Executive Summary*, MR-710-AF, RAND, Santa Monica, Calif.

Ho, Yu-Chi (ed.). 1989. *Special Issue on Discrete Event Dynamic Systems of the Proceedings of the IEEE*, 77(1), January. Also found in Ho, Yu-Chi (ed.). *Discrete Event Dynamic Systems: Analyzing Complexity and Performance in the Modern World, IEEE*, Piscataway, N.J., 1992.

Hodges, James, and James Dewar. 1992. *Is It You or Your Model Talking? A Framework for Model Validation*, R-4114-AF/A/OSD, RAND, Santa Monica, Calif.

Hoffmann, Hans W. 1989. *Einsatz von Expertensystemtechniken in Gefechtssimulationen*, Oberpfaffenhofen, Carl Cranz Gesellschaft, FRG.

Holland, John. 1995. *Hidden Order*, Addison-Wesley, Reading, Mass.

Hong, G.P., and T.G. Kim. 1996. "A Framework for Verifying Discrete Event Models Using a Dual Specification Approach," *Trans. Soc. Comput. Simul.*, 13(1):19-34.

Horrigan, Timothy J. 1991. "Configural Theory and the Mathematical Modeling of Combat," Horrigan Analytics, HAS 91-17-1, Chicago, Ill.

Horrigan, Timothy J. 1993. "The 'Configuration Problem' and Challenges for Aggregation" in Davis, Paul K., and Richard Hillestad. 1992. "Using Simulation in the Education of General Officers," *Proceedings of the 1992 Summer Computer Simulation Conference*, Society for Computer Simulation, San Diego, Calif., pp. 1071-1076.

Horrigan, Timothy J. 1995. "Configuration and the Effectiveness of Air Defense Systems in Simplified, Idealized Combat Situations—A Preliminary Examination," Horrigan Analytics, Chicago, Ill., June.

Huber, Reiner K. (ed.). 1990. "An Analytical Approach to Cooperative Security: First Order Assessment of Force Posture Requirements in a Changing Security Environment," *Military Stability: Prerequisites and Analysis Requirements for Conventional Stability in Europe, Nomos* Verlagsgesellschaft, Baden-Baden.

Huber, Reiner K. 1992. "Re: McQuie on the 3:1 Rule," *Phalanx*, 25(1):28-29.

Hughes, Wayne P. 1986. *Fleet Tactics: Theory and Practice,* Naval Institute Press, Annapolis, Md.

Hughes, Wayne P. (ed.). 1989. *Military Modeling*, rev. ed., Military Operations Research Society, Alexandria, Va.

Ilachinski, Andrew. 1996a. *Land Warfare and Complexity, Part 1: Mathematical Background and Technical Sourcebook*, Research memorandum 96-68 and CIM461, Center for Naval Analyses, Alexandria, Va., July.

Ilachinski, Andrew. 1996b. *Land Warfare and Complexity, Part 2: An Assessment of the Applicability of Non-Linear Dynamics and Complex Systems Theory to the Study of Land Warfare*, Research memorandum 96-68, Center for Naval Analyses, Alexandria, Va., July.

Innis, G., and E. Rexstad. 1983. "Simulation Model Simplification Techniques," *Simulation,* July:7-15.

Jefferson, Lt Col Mark, USAF. 1996. Defense Modeling and Simulation Office, Office of the Under Secretary of Defense for Acquisition and Technology, Washington, D.C., April 2.

Johnson, Jay, ADM, USN, Chief of Naval Operations. 1997. Address to the U.S. Naval Institute, Annapolis, Md., April 23.

Johnson, Stuart, and Martin C. Libicki (eds.). 1996. *Dominant Battlespace Knowledge*, National Defense University, Institute for National Strategic Studies, Washington, D.C.

Kaminski, Paul G., Under Secretary of Defense for Acquisition and Technology. 1996. "DMSO 'Modeling and Simulation,'" Keynote address at DOD Fifth Annual Industry Briefing, Alexandria, Va., May 22.

Kim, J., and B.P. Zeigler. 1996. "Designing Fuzzy Logic Controllers Using a Multi-resolutional Search Paradigm," *IEEE Transactions on Fuzzy Systems*, 4(3):213-226.

Kuck, Inara. 1996. *Warfare Simulation: Status and Issues for Space, Final Report*, Phillips Laboratory, U.S. Air Force Materiel Command, Kirtland AFB, N.Mex.

Lanchester, Fred W. 1916. *Aircraft in Warfare: The Dawn of the Fourth Arm*, Constable, London.

Lempert, Robert, Michael E. Schlesinger, and Steve C. Bankes. 1996. *When We Don't Know the Costs or the Benefits: Adaptive Strategies for Abating Climate Change,* RAND, Santa Monica, Calif. Reprint from the journal article in *Climatic Change*, 33(2), June 1996.

Lewis, Robert. 1997. "Designing Verification, Validation, and Accreditation to Cost," speech given at the conference, *Linking Simulations for Analysis,* Military Operations Research Society (MORS), Albuquerque, N.Mex., February 25-26, 1997. July 15. See also Lewis, Robert O. 1992. *Independent Verification and Validation: A Life Cycle Engineering Process for Quality Software* (New Dimensions in Engineering), John Wiley and Sons, New York, October.

Louri, Ahmed, Hongki Sung, Yoonkeon Moon, and Bernard P. Zeigler. 1995. "An Efficient Signal Distinction Scheme for Large-scale Free-space Optical Networks Using Genetic Algorithms," *Photonics in Switching: Topical Meeting*, The Optical Society of America, Salt Lake City, Utah, pp. 90-92.

Lucas, Tom, Bob Kerchner, John Friel, and Dan Jones. 1997. *A Guide for Analysis Using Advanced Distributed Simulation (ADS),* draft, RAND, Santa Monica, Calif., January.

MacGregor, Douglas A. 1997. *Breaking the Phalanx, A New Design for Landpower in the 21st Century*, Praeger, Westport, Conn.

Matsumura, John, and Randall Steeb. 1996. *Analytical Support to the DSB*, DRR-1501-AF, RAND, Santa Monica, Calif., August.

Mattock, Michael G., John Schank, James Stucker, and Jeff Rothenberg. 1995. *New Capabilities for Strategic Mobility Analysis Using Mathematical Programming,* MR-296-JS, RAND, Santa Monica, Calif.

Mearsheimer, John J. 1989. "Assessing the Conventional Balance: The 3:1 Rule and Its Critics," *International Security,* 13(4):54-89, Spring.

Miachalewicz, Z. 1992. *Genetic Algorithm + Data Structure = Evolution Programming.* Springer-Verlag, New York.

Military Operations Research Society (MORS). 1992. *Simulation Validation Workshop Proceedings (SIMVAL II), MORS,* Adele Ritchie (ed.), chapters by Marion Williams and James Sikora; Joseph Cynamon; Jim Metzger; Donald Giadrosich; Dale Henderson; Don Gaver; Ernest Seglie and Patricia Sanders; Paul K. Davis, MORS, Alexandria, Va.

Military Operations Research Society (MORS). 1997. *Linking Simulations for Analysis,* held at Albuquerque, N.Mex., February 25-26.

Miller, Duncan C., and Jack A. Thorpe. 1995. "SIMNET: The Advent of Simulator Networking," *IEEE Proceedings,* 83(8):1114-1123, August.

Moon, Y. 1996. *High Performance Simulation Based Optimization Environment for Large Scale Systems,* University of Arizona, Tucson, Ariz.

Morgeson, D. 1996. *High Level Architecture (HLA) Implementation Issues,* Los Alamos National Laboratory, Los Alamos, N.Mex.

Morrison, John D. 1996. "Realtime Learning of Doctrine and Tactics Using Neural Networks and Combat Simulations," *Military Operations Research,* 2(3):45-60.

Muessig, Paul. 1997. "Surprise, Surprise! VV&A Yields a Few Unexpected Results," *Phalanx,* 30(2):6, June.

Natrajan, Anand and Anh Nguyen-Tuong. 1995. To Disaggregate Or Not To Disaggregate, That Is *Not* the Question, electronic paper for the *Electronic Conference on Scalability in Training Simulation* (*ELECSIM 1995*). (Web site: http://www.mystech.com/~smithr/elecsim95/)

Naval Research Advisory Committee. 1994. *Report on Modeling and Simulation,* Office of the Assistant Secretary of the Navy for Research, Development, and Acquisition, Arlington, Va.

Naval Studies Board. 1996. *The Navy and Marine Corps in Regional Conflict in the 21st Century,* National Academy Press, Washington, D.C.

Neuberger, Thomas K., and Dennis P. Shea. 1995. *Applying Synthetic Environments to Operational Training: A Perspective from Kernel Blitz 1995,* Center for Naval Analyses, Alexandria, Va.

Office of Technology Assessment. 1995. *Distributed Interactive Simulation of Combat,* U.S. Congress, Washington, D.C.

Olsen, Paul. 1976. *A Maneuver Oriented Model of Conventional Land Warfare: IDAHEX,* Institute for Defense Analyses, Arlington, Va.

O'Neil, William D. 1997. "Campaign Analysis and Modeling: A Review and Comparative Evaluation," draft, Center for Naval Analyses, March.

Palmore, Julian. 1993. "Variable Resolution Modeling in Mathematics," *Proceedings of Conference on Variable Resolution Modeling,* Davis and Hillestad (eds.), RAND, CF-103, Santa Monica, Calif.

Patenaude, Anne. 1996. *Study on Effectiveness of Modeling and Simulation in Weapons Systems Acquisition Process,* (study conducted for the Deputy Director, Test, Systems Engineering and Evaluation, Office of the Secretary of Defense), Science Applications International Corporation, McLean, Va., October.

Pettey, Chrisila B., Michael R. Lueze, and John J. Grefenstette. 1987. "A Parallel Genetic Algorithm," *Proceeding of the 2nd International Conference on Genetic Algorithms and Their Applications,* Lawrence Erlbaum Associates, Hillsdale, N.J., July, pp. 155-161.

Pichler, F., and H. Schwartzel (eds.). 1992. *CAST (Computer Aided System Theory) Methods in Modeling,* Springer-Verlag, Berlin.

Praehofer, H. 1991. *System Theoretic Foundations for Combined Discrete-Continuous System Simulation,* Johannes Kepler University of Linz, Linz, Austria.

Quade, Edward S., and Grace M. Carter (eds.). 1989. *Analysis for Public Decisions*, 3d ed., Prentice Hall, Old Tappan, N.J.

Rothenberg, J. 1986. "Object-oriented Simulation: Where Do We Go From Here?" *Proceedings of the 1986 Winter Simulation Conference,* Washington, D.C., Dec. 8-10, J. Wilson, J. Henriksen, and S. Roberts (eds.), The Society for Computer Simulation, San Diego, Calif., pp. 464-469. Reprinted as N-3028-DARPA, RAND, Santa Monica, Calif., October.

Rothenberg, Jeff. 1989. "The Nature of Modeling," *Artificial Intelligence, Simulation, and Modeling*, Sidman, Loparo, and Nielson (eds.), John Wiley and Sons, New York, pp. 75-92. Reprinted as 3207-DARPA, RAND, Santa Monica, Calif.

Rowland, D., L.R. Speight, and M.C. Keys. 1996. "Maneuver Warfare: Some Conditions Associated with Success at the Operational Level," *Military Operations Research,* 2(3):5-16.

Ruelle, David. 1991. *Chance and Chaos*, Princeton Science Library, Princeton, N.J.

Rumbaugh, James, Michael Blaha, William Premerlani, James Edy, and William Lorensen. 1991. *Object-Oriented Modeling and Design*, Prentice-Hall, Englewood Cliffs, N.J.

Sanders, Patricia, and Ray Miller. 1995. "Model Verification, Validation, and Accreditation (VV&A): Common Ground Within the M&S Community," a white paper based on a 1995 colloquium at Ft. Belvoir, Va. Available from DMSO's document library at Web site: http://www.dmso.mil/docslib/mspolicy/msmp/1095msmp/.

Sarjoughian, Hessam S. 1995. *Inductive Modeling of Discrete-event Systems: A TMS-based Nonmonotonic Reasoning Approach,* Ph.D. thesis, Department of Electrical and Computer Engineering, University of Arizona, Tucson, Ariz. Web site: http://hpcc.ece.arizona.edu/~hessam/ohtml/phd-95.html.

Scales, John Richard. 1995. "A Modified Lanchester Linear Process Calibrated to Historical Data," *Naval Research Logistics,* 42(3):491-502, April. Also in Bracken et al. (1995), pp. 345-356.

Scheckeler, Günther. 1993. "A Hierarchy of Models at Different Resolution Levels," in Davis and Hillestad (1993).

Sevinc, Suleyman. 1990. "Automation of Simplification in Discrete Event Modeling and Simulation," *International Journal of General Systems,* 18:125-145.

Shalikashvili, John, GEN, USA. 1996. *Joint Vision 2010,* Joint Chiefs of Staff, Washington, D.C.

Shiflett, James E., W.H. (Dell) Lunceford, and Ruth P. Willis. 1995. "Application of Distributed Interactive Simulation Technology Within the Department of Defense," *Proceedings of the IEEE,* 83(8):1168-1178, August.

Sikora, Jim, and Marion Williams. 1997. "Verification, Validation, and Accreditation—What Is It Today," *Phalanx*, 30(2):1, June.

Simard, R. 1996. *Simulation EnVironment: SimVir, A Simulation Development Process* in *AI, Simulation and Planning in High Autonomy Systems,* San Diego, Calif. Available at Web site: http://www-ais.ece.arizona.edu.

Simon, Herbert A. 1990. "Prediction and Prescription in Systems Modeling," *Operations Research,* 38(1):7-14, January-February.

Simon, Herbert A. 1996. *Sciences of the Artificial,* 3d ed., Massachusetts Institute of Technology Press, Cambridge, Mass.

Simpson, H., W.D. West, and D. Gleisner. 1995. *The Use of Simulation in Military Training: Value, Investment, and Potential,* DMDC TR 95-007, Defense Manpower Data Center, Monterey, Calif.

Stanley, Walter. 1997. "The Verification, Validation, and Certification of Data in Modeling and Simulation Applications," *Phalanx,* 30(2):12, June.

Taylor, James G. 1983a. "An Introduction to Lanchester-Type Models of Warfare," in *Proceedings of the Workshop on Modeling and Simulation of Land Combat,* L.G. Callahan (ed.), Department of Continuing Education, Georgia Institute of Technology, Atlanta, Ga.

Taylor, James G. 1983b. *Lanchester Models of Warfare*, two volumes, Operations Research Society of America, Arlington, Va.

U.S. Air Force. 1996. *Verification, Validation, and Accreditation (VV&A)*, HQ USAF/XOMT, Technical Support Division, Directorate of Modeling, Simulation, and Analysis, Washington, D.C., June 1. (Web site: http://www.fas.org/irp/doddir/usaf/16-1001.htm)

U.S. Air Force Scientific Advisory Board. 1995. *New World Vistas: Air and Space Power for the 21st Century*,.Air Force Scientific Advisory Board, Washington, D.C.

U.S. Army Training and Doctrine Command (TRADOC). 1994. *Force XXI Operations, A Concept for the Evolution of Full-Dimensional Operations for the Strategic Army of the Early Twenty-First Century*, U.S. Army, Fort Monroe, Va., August 1. (Web site: http://www-tradoc.monroe.army.mil/tpubs/pams/p525-5toc.htm)

U.S. Department of Defense (DOD). 1995. *DOD FY 95 Master Plan for Target Interaction, Lethality, and Vulnerability (TILV) Science and Technology (S&T) Programs*, Volume I, Classical Ballistic Threats, May 4 (Revised).

U.S. General Accounting Office. 1995. *Bottom-Up Review: Analysis of Key DOD Assumptions*, NSIAD-95-56, Washington, D.C., January 31.

U.S. Navy. 1996. *Doctrine for Command and Control*, Department of the Navy, Washington, D.C.

Vahie, S. and N. Jouppi. 1996. "Dynamic Neuronal Ensembles: A New Paradigm for Learning & Control," *Annual Conference on AI, Simulation and Planning in High Autonomy Systems*, San Diego, Calif. Available at Web site: www-ais.arizona.edu.

Vector Research, Inc. 1992. *Global Force Capability Requirements for U.S. Defense Planning*, two volumes, VRI-TRADOC-13.17, FR92-1, Ann Arbor, Mich., October.

Wainstein, Leonard. 1986. *The Relationship of Battle Damage to Unit Combat Performance*, Institute for Defense Analyses, Alexandria, Va.

Wise, Ben. 1991. "Lanchester Equations for Automated Planning," *Phalanx*, 24(1):24-27, 36.

Worley, D. Robert, Henry Simpson, Franklin Moses, Matt Aylward, Michelle Bailey, and Dean Fish. 1996. *Utility of Modeling and Simulation in the Department of Defense: Initial Data Collection*, Institute for Defense Analyses, Alexandria, Va., August.

Young, Michael J. 1992. *Successively Approximating Human Performance*, AL/HR-TP-1993-0026, Armstrong Laboratory, Logistics Research Division, Wright-Patterson AFB, Ohio.

Youngblood, Simone. 1997. "The Joint Warfare System (JWARS) VV&A Effort: Challenges in VV&A of a Joint Object-Oriented Development," *Phalanx*, 30(2):11, June.

Youngblood, Simone, and Dale Pace. 1993, *Preliminary Bibliography for VV&A*, NWA-93-036, Applied Physics Laboratory, Johns Hopkins University, Laurel, Md. See also the final report, Simone Youngblood, *Bibliography for Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)*, NWA-93-037, Applied Physics Laboratory, Johns Hopkins University, Laurel, Md., August, 1993.

Zadeh, Lofti. 1994. "Soft Computing and Fuzzy Logic," *IEEE Software*, 11(6):48-56, November.

Zeigler, Bernard. 1976. *Theory of Modeling and Simulation*, John Wiley, New York.

Zeigler, Bernard P. 1979. "Structuring Principles for Multifaceted System Modelling," in *Methodology in Systems Modelling and Simulation*, B.P. Zeigler, M. Elzas, et al. (eds.), North Holland Press, New York, pp. 93-138.

Zeigler, Bernard. 1984. *Multifaceted Modeling and Discrete Event Simulation*, Academic Press, London.

Zeigler, Bernard. 1989. "DEVS Representation of Dynamical Systems: Event-based Intelligent Control" *Proceedings of IEEE*, 77(1):72, January.

Zeigler, Bernard. 1990. *Object-oriented Simulation with Hierarchical, Modular Models: Intelligent Agents and Endomorphic Systems*, Academic Press, Boston, Mass., and San Diego, Calif.

Zeigler, Bernard, H. Praehofer, and Jerry Rozenblit. 1993. "Integrating System Formalisms: How Object Orientation Supports CAST for Intelligence Systems Design," *Journal of Systems Engineering*, 3:209-219.

Zeigler, Bernard, Y. Moon, D. Kim, and G.J. Ball. 1997. "The DEVS Environment for High Performance Modeling and Simulation," *IEEE Computational Science and Engineering*, 4(3), July-September, in press.

APPENDICES

A

Terms of Reference

CHIEF OF NAVAL OPERATIONS

28 November 1995

Dear Dr. Alberts,

In 1986, at the request of this office, the Academy's Naval Studies Board undertook a study entitled "Implications of Advancing Technology for Naval Warfare in the Twenty-First Century." The Navy-21 report, as it came to be called, projected the impact of evolving technologies on naval warfare out to the year 2035, and has been of significant value to naval planning over the intervening years. However, as was generally agreed at the time, the Navy and Marine Corps would derive maximum benefit from a periodic comprehensive review of the implications of advancing technology on future Navy and Marine Corps capabilities. In other words, at intervals of about ten years, the findings should be adjusted for unanticipated changes in technology, naval strategy, or national security requirements. In view of the momentous changes that have since taken place, particularly with national security requirements in the aftermath of the Cold War, I request that the Naval Studies Board immediately undertake a major review and revision of the earlier Navy-21 findings.

The attached Terms of Reference, developed in consultation between my staff and the Chairman and Director of the Naval Studies Board, indicate those topics which I believe should receive special attention. If you agree to accept this request, I would appreciate the results of the effort in 18 months.

Sincerely,

J. M. BOORDA
Admiral, U.S. Navy

Dr. Bruce M. Alberts
President
National Academy of Sciences
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

NATIONAL ACADEMY OF SCIENCES
DEC 01 1995
OFFICE OF THE PRESIDENT

Enclosure

TERMS OF REFERENCE

TECHNOLOGY FOR THE FUTURE NAVY

The Navy-21 study (Implications of Advancing Technology for Naval Warfare in the Twenty-First Century), initiated in 1986 and published in 1988, projected the impact of technology on the form and capability of the Navy to the year 2035. In view of the fundamental national and international changes -- especially the Cold War's end -- that have occurred since 1988, it is timely to conduct a comprehensive review of the Navy-21 findings, and recast them, where needed, to reflect known and anticipated changes in the threat, naval missions, force levels, budget, manpower, as well as present or anticipated technical developments capable of providing cost effective leverage in an austere environment. Drawing upon its subsequent studies where appropriate, including the subpanel review in 1992 of the prior Navy-21 study, the Naval Studies Board is requested to undertake immediately a comprehensive review and update of its 1988 findings. In addition to identifying present and emerging technologies that relate to the full breadth of Navy and Marine Corps mission capabilities, specific attention also will be directed to reviewing and projecting developments and needs related to the following: (1) information warfare, electronic warfare, and the use of surveillance assets; (2) mine warfare and submarine warfare; (3) Navy and Marine Corps weaponry in the context of effectiveness on target; (4) issues in caring for and maximizing effectiveness of Navy and Marine Corps human resources. Specific attention should be directed, but not confined to, the following issues:

1. Recognizing the need to obtain maximum leverage from Navy and Marine Corps capital assets within existing and planned budgets, the review should place emphasis on surveying present and emerging technical opportunities to advance Navy and Marine Corps capabilities within these constraints. The review should include key military and civilian technologies that can affect Navy and Marine Corps future operations. This technical assessment should evaluate which science and technology research must be maintained in naval research laboratories as core requirements versus what research commercial industry can be relied upon to develop.

2. Information warfare, electronic warfare and the exploitation of surveillance assets, both through military and commercial developments, should receive special attention in the

review. The efforts should concentrate on information warfare, especially defensive measures that affordably provide the best capability.

3. Mine warfare and submarine warfare are two serious threats to future naval missions that can be anticipated with confidence, and should be treated accordingly in the review. This should include both new considerations, such as increased emphasis on shallow water operations, and current and future problems resident in projected worldwide undersea capability.

4. Technologies that may advance cruise and tactical ballistic missile defense and offensive capabilities beyond current system approaches should be examined. Counters to conventional, bacteriological, chemical and nuclear warheads should receive special attention.

5. The full range of Navy and Marine Corps weaponry should be reviewed in the light of new technologies to generate new and improved capabilities (for example, improved targeting and target recognition).

6. Navy and Marine Corps platforms, including propulsion systems, should be evaluated for suitability to future missions and operating environments. For example, compliance with environmental issues is becoming increasingly expensive for the naval service and affects operations. The review should take known issues into account, and anticipate those likely to affect the Navy and Marine Corps in the future.

7. In the future, Navy and Marine Corps personnel may be called upon to serve in non-traditional environments, and face new types of threats. Application of new technologies to the Navy's medical and health care delivery systems should be assessed with these factors,.as well as joint and coalition operations, reduced force and manpower levels, and the adequacy of specialized training in mind.

8. Efficient and effective use of personnel will be of critical importance. The impact of new technologies on personnel issues, such as education and training, recruitment, retention and motivation, and the efficient marriage of personnel and machines should be addressed in the review. A review of past practices in education and training would provide a useful adjunct.

9. Housing, barracks, MWR facilities, commissaries, child care, etc. are all part of the Quality of Life (QOL) of naval personnel. The study should evaluate how technology can be used to enhance QOL and should define militarily meaningful measures of effectiveness (for example, the impact on Navy readiness).

10. The naval service is increasingly dependent upon modeling and simulation. The study should review the overall architecture of models and simulation in the DoD (DoN, JCS, and OSD), the ability of models to represent real world situations, and their merits as tools upon which to make technical and force composition decisions.

The study should take 18 months and produce a single-volume overview report supported by task group reports (published either separately or as a single volume). Task group reports should be published as soon as completed to facilitate incorporation into the DoN planning and programming process. An overview briefing also should be produced that summarizes the contents of the overview report, including the major findings, conclusions, and recommendations.

B

Virtual Engineering: Toward a Theory for Modeling and Simulation of Complex Systems

John Doyle, California Institute of Technology

INTRODUCTION

This paper is a primer surveying a wide range of issues tied together loosely in a problem domain tentatively referred to as "virtual engineering" (VE). This domain is concerned with modeling and simulation of uncertain, heterogeneous, complex, dynamical systems—the very kind of M&S on which much of the vision discussed in this study depends. Although the discussion is wide ranging and concerned primarily with topics distant from those usually discussed by the Department of the Navy and DOD modeling communities, understanding how those topics relate to one another is essential for appreciating both the potential and the enormous intellectual challenges associated with advanced modeling and simulation in the decades ahead.

BACKGROUND

Perhaps the most generic trend in technology is the creation of increasingly complex systems together with a greater reliance on simulation for their design and analysis. Large networks of computers with shared databases and high-speed communication are used in the design and manufacture of everything from microchips to vehicles such as the Boeing 777. Advances in technology

NOTE: This appendix benefited from material obtained from many people and sources: Gabriel Robins on software, VLSI, and the philosophy of modeling, Will O'Neil on CFD, and many colleagues and students.

have put us in the interesting position of being limited less by our inability to sense and actuate, to compute and communicate, and to fabricate and manufacture new materials, than by how well we understand, design, and control their interconnection and the resulting complexity. While component-level problems will continue to be important, systems-level problems will be even more so. Further, "components" (e.g., sensors) increasingly need to be viewed as complex systems in their own right. This "system of systems" view is coming to dominate technology at every level. It is, for example, a basic element of DOD's thinking in contexts involving the search for dominant battlefield awareness (DBA), dominant battlefield knowledge (DBK), and long-range precision strike.

At the same time, virtual reality (VR) interfaces, integrated databases, paperless and simulation-based design, virtual prototyping, distributed interactive simulation, synthetic environments, and simultaneous process/product design promise to take complex systems from concept to design. The potential of this still-nascent approach is well appreciated in the engineering and science communities, but what "it" is is not. For want of a better phrase, we refer to the general approach here as "virtual engineering" (VE). VE focuses on the role of M&S in *uncertain, heterogeneous, complex, dynamical systems*—as distinct from the more conventional applications of M&S. But VE, like M&S, should be viewed as a problem domain, not a solution method.

In this paper, we argue that the enormous potential of the VE vision will not be achieved without a sound theoretical and scientific basis that does not now exist. In considering how to construct such a base, we observe a unifying theme in VE: Complexity is a by-product of designing for *reliable predictability in the presence of uncertainty* and subject to resource limitations.

A familiar example is smart weapons, where sensors, actuators, and computers are added to counter uncertainties in atmospheric conditions, release conditions, and target movement. Thus, we add complexity (more components, each with increasing sophistication) to reduce uncertainties. But because the components must be built, tested, and then connected, we are introducing not only the potential for great benefits, but also the potential for *catastrophic failures* in programs and systems. Evaluating these complexity versus controllability tradeoffs is therefore very important, but also can become conceptually and computationally overwhelming.

Because of the critical role VE will play, this technology should be robust, and its strengths and limitations must be clearly understood. The goal of this paper is to discuss the basic technical issues underlying VE in a way accessible to diverse communities—ranging from scientists to policy makers and military commanders. The challenges in doing so are intrinsically difficult issues, intensely mathematical concepts, an incoherent theoretical base, and misleading popular expositions about "complexity."

APPROACH

In this primer on VE, we concentrate on "physics-based" complex systems, but most of the issues apply to other M&S areas as well, including those involving "intelligent agents." Our focus keeps us on a firmer theoretical and empirical basis and makes it easier to distinguish the effects of complexity and uncertainty from those of simple lack of knowledge. Our discussion also departs from the common tendency to discuss VE as though it were a mere extension of software engineering. Indeed, we argue *that uncertainty management in the presence of resource limitations* is the dominant technical issue in VE, that conventional methods for M&S and analysis will be inadequate for large complex systems, and that VE requires new mathematical and computational methods (VE theory, or VET). We need a more integrated and coherent theory of modeling, analysis, simulation, testing, and model identification from data, and we must address nonlinear, interconnected, heterogeneous systems with hierarchical, multi-resolution, variable-granularity models—both theoretically and with suitable software architectures and engineering environments.

Although the foundations of any VE theory will be intensely mathematical, we rely here on concrete examples to convey key ideas. We start with simple physical experiments that can be done easily with coins and paper to illustrate dynamical systems concepts such as sensitivity to initial conditions, bifurcation, and chaos. We also use these examples to introduce uncertainty modeling and management.

Having introduced key ideas, we then review major success stories of what could be called "proto-VE" in the computer-aided design (CAD) of the Boeing 777, computational fluid dynamics (CFD), and very large scale integrated circuits (VLSI). While these success stories are certainly encouraging, great caution should be used in extrapolating to more general situations. Indeed, we should all be sobered by the number of major failures that have already occurred in complex engineering systems such as the Titanic, Tacoma-Narrows bridge, Denver baggage-handling system, and Ariane booster. We argue that *uncertainty management* together with *dynamics and interconnection* is the key to understanding both these successes and failures and the future challenges.

We then discuss briefly significant lessons from software engineering and computational complexity theory. There are important generalizable lessons, but—as we point out repeatedly—software engineering is not a prototype for VE. Indeed, the emphasis on software engineering to the exclusion of other subjects has left us in a virtual "pre-Copernican" stage in important areas having more to do with the *content* of M&S for complex systems.

Against this background, we draw implications for VE. We go on to relate these implications to famous failures of complex engineering systems, thereby demonstrating that the issues we raise are not mere abstractions, and that achieving the potential of VE (and M&S) will be enormously challenging. We touch

briefly on current examples of complex systems (smart weapons and airbags) to relate discussion to the present. We then discuss what can be learned from control theory and its evolution as we move toward a theory of VE. At that point, we return briefly to the case studies to view them from the perspective of that emerging theory. Finally, we include a section on what we call "soft computing," a domain that includes "complex-adaptive-systems research," fuzzy logic, and a number of other topics on which there has been considerable semi-popular exposition. Our purpose is to relate these topics to the broader subject of VE and to provide readers with some sense of what can be accomplished with "soft computing" and where other approaches will prove essential.

In summary before getting into our primer, we note that several trends in M&S of complex systems are widely appreciated, if not well understood. There is an increasing emphasis on moving problems and models from linear to nonlinear; from static to dynamic; and from isolated and homogeneous to heterogeneous, interconnected, hierarchical, and multi-resolution (or variable granularity and fidelity). What is poorly understood is the role of uncertainty, which we claim is actually the origin of all the other trends. Model uncertainty arises from the differences between the idealized behavior of conventional models and the reality they are intended to represent. The need to produce models that give reliable predictability of complex phenomena, and thus have limited uncertainty, leads to the explicit introduction of dynamics, nonlinearity, and hierarchical interconnections of heterogeneous components. Thus the focus of this paper is that uncertainty is the key to understanding complex systems.

INTRODUCTION TO CENTRAL CONCEPTS

Dynamical Systems

A few simple thought experiments can illustrate the issues of uncertainty and predictability—as well as of nonlinearity, dynamics, heterogeneity, and ultimately complexity. Most of the experiments we discuss here can also be done with ordinary items like coins and paper.

Consider a coin-tossing mechanism that imparts a certain linear and angular velocity on a coin, which is then allowed to bounce on a large flat floor, as depicted in Figure B.1. Without knowing much about the mechanism, we can reliably predict that the coin will come to rest on the floor. For most mechanisms, it will be impossible to predict whether it will be heads or tails. Indeed, heads or tails will be equally likely, and any sequence of heads or tails will be equally likely. Such specific predictions are as reliably *unpredictable* as the eventual stopping of the coin is predictable.

The reliable unpredictability of heads or tails is a simple consequence of the sensitivity to initial conditions that is almost inevitable in such a mechanism. The coin will bounce around on the floor in an apparently random and erratic manner

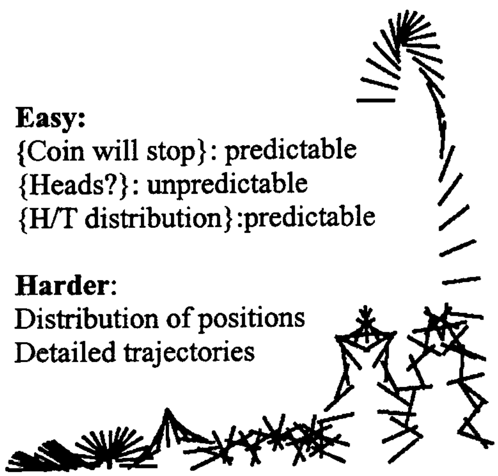

FIGURE B.1 Coin tossing experiment.

before eventually coming to rest on the floor. The coin's trajectory will be different in detail for each different toss, in spite of efforts to make the experiment repeatable. Extraordinary measures would be needed to ensure predictability (e.g., dropping the coin heads up a short distance onto a soft and sticky surface, so as always to produce heads).

Sensitivity to initial conditions (STIC) can occur even in simple settings such as a rigid coin in a vacuum with no external forces, not even gravity. With zero initial velocity, the coin will remain stationary, but the smallest initial nonzero velocity will cause the coin to drift away with distance proportional to time. The dynamics are linear and trivial. This points out that—in contrast with what is often asserted—sensitivity to initial conditions is very much a linear phenomenon. Moreover, even in nonlinear systems, the standard definition of sensitivity involves examining infinitesimal variations about a given trajectory and examining the resulting *linear* system. Thus even in nonlinear systems, sensitivity to initial conditions boils down to the behavior of linear systems. What nonlinearity contributes is making it more difficult to completely characterize the consequences of sensitivity to initial conditions.

Sensitivity to initial conditions is also a matter of degree; the coin-in-free-space example being on the boundary of systems that are sensitive to initial conditions. Errors in initial conditions of the coin lead to a drifting of the trajectories that grows linearly with time. In general, the growth can be exponential, which is more dramatic. If we add atmosphere, but no other external force, the coin will eventually come to rest no matter what the initial velocities, so this

is clearly less sensitive to initial conditions than the case with no atmosphere. A coin in a thick, sticky fluid like molasses is even less sensitive.

Not all features of our experiment are sensitive to initial conditions. The final vertical position is reliably predictable, the time at which the coin will come to rest is less so, the horizontal resting location even less so, and so on, with the heads or tails outcome perfectly unpredictable. It follows that any notion of complexity cannot be attributed to the system, but must include the property of it that is in question.

EXPONENTIAL GROWTH, CHAOS, AND BIFURCATION

We can get a better understanding of sensitivity to initial conditions with some elementary mathematics. Suppose we have a model of the form $x(t+1) = f(x(t))$. This tells us what the state variable x is at time $t+1$ as a function of the state x at time t. This is called a difference equation, which is one way to describe a dynamical system—i.e., one that evolves with time. If we specify $x(t)$ at some time, say $t = 0$, then the formula $x(t+1) = f(x(t))$ can be applied recursively to determine $x(t)$ for all future times $t = 1,2,3, \ldots$. This determines an orbit or trajectory of the dynamical system. This only gives x at discrete times, and x is undefined elsewhere. It is perhaps more natural to model the coin and other physical systems with differential equations that specify the state at all times, but difference equations are simpler to understand. For the coin, the state would include at least the positions and velocities of the coin, and possibly some variables to describe the time evolution of the air around the coin. If the coin were flexible, the state might include some description of the bending and its rate. And so on.

A scalar linear difference equation is of the form $x(t+1) = ax(t)$, where a is a constant (the vector case is $x(t+1) = Ax(t)$, where A is a matrix). If $x(0)$ is given, the solution for all time is $x(t) = a^t x(0)$. Thus, if $a>1$, nonzero solutions grow exponentially and the system is called unstable. Since the system is linear, any difference in initial conditions will also grow exponentially. (If $a<1$, then solutions decay exponentially to zero and the origin is a stable fixed point.)

Exponential growth appears in so many circumstances that it is worth dramatizing its consequences. If $a = 10$, then in each second x gets 10 times larger, and after 100 seconds it is 10^{100} larger. With this type of exponential growth, an error smaller than the nucleus of a hydrogen atom would be larger than the diameter of the known universe in less than 100 seconds. Of course, no physical system could have this as a reasonable model for long time periods. The point is that linear systems can exhibit very extreme sensitivity to initial conditions because of exponential growth. Of course, STIC is a matter of degree. The quantity $\ln(a)$ is one measure of the degree of STIC and is called the Lyapunov exponent.

Suppose we modify our scalar linear system slightly to make it the nonlinear system $x(t+1) = 10x(t) \bmod 10$ and restrict the state to the interval $[0,10]$. This system can be thought of as taking the decimal expansion of $x(t)$ and shifting the

decimal point to the right and then truncating the digit to the left of the units place. For example, if $x(0) = \pi = 3.141592$ ..., then $x(1) = 1.41592$ ... and $x(2) = 4.1592$ ... and so on. This still has, in the small, the same exponential growth as the linear system, but its orbits stay bounded. If $x(0)$ is rational, then the $x(t)$ will be periodic, and thus there are a countable number of periodic orbits (arbitrarily long periods). If $x(0)$ is irrational, then the orbit will stay irrational and not be periodic, but it will appear exactly as random and irregular as the irrational initial condition. As is well-known, this system exhibits deterministic chaos. The Lyapunov exponent can also be generalized to nonlinear systems, and in this case would still be $\ln(a)$.

The several alternative mathematical definitions of chaos are all beyond the scope of this paper, but the essential features of chaotic systems are sensitivity to initial conditions (STIC), periodic orbits with arbitrarily long periods, and an uncountable set of bounded nonperiodic (and apparently random) orbits.

The STIC property and the large number of periodic orbits can occur in linear systems. But the "arbitrarily long periods" and "bounded, nonperiodic, apparently random" features require some nonlinearity. Chaos has received much attention in the popular press, which often confuses nonlinearity and sensitivity to initial conditions in suggesting that the former in some way causes the latter, when in fact both are independent and necessary but not sufficient to create chaos. The formal mathematical definitions of chaos involve infinite time horizon orbits, so none of our examples so far would be, strictly speaking, chaotic. A simple way to get a system that is closer in spirit to chaos would be to put our coin in a box and then shake the box with some periodic motion. Even though the box had regular motion, under many circumstances the coin's motion in bouncing around the box would appear random and irregular.

A simple model with linear dynamics between collisions and a linear model for the collisions with the box would almost certainly be chaotic, although even this simple system is too complicated to prove the existence of chaos rigorously and it must be suggested via simulation. Very few dynamical systems have been proved chaotic, and most models of physical systems that appear to exhibit chaos are only suggested to be so by simulation. One-degree-of-freedom models of a ball in a cylinder with a closed top and a periodically moving piston have been proved chaotic. The ball typically bounces between the piston and the other end wall of the cylinder with the impact times being random, even though the dynamics are purely deterministic, and even piecewise linear.

To get a sense of the notion of bifurcation in dynamical systems, consider the following experiment. Drop a quarter in as close to a horizontal position and with as little initial velocity as possible. It will drop nearly straight down, and the air will have little effect at the speeds the coin will attain while it bounces around the floor. Now take a quarter-size piece of paper and repeat the experiment. The paper will begin fluttering rapidly and fall toward the floor at a large angle, landing far away from where a real quarter would have first hit the floor. This is

an example of a bifurcation, where a seemingly small change in properties creates a dramatic change in behavior. The heavy coin will reliably and predictably hit the floor beneath where it is dropped (at which point subsequent collisions may make what follows it quite unpredictable), whereas the paper coin will spin off in any direction and land far away, but then quickly settle down without bouncing. Thus one exhibits STIC, while the other does not.

A simple variant on this experiment illustrates a bifurcation more directly. Make two photocopies of the diagram in Figure B.2 (or just fold pieces of paper as follows), and cut out the squares along the solid line. The unfolded paper will flutter erratically when dropped, exhibiting STIC. Next, take one of the papers and fold it along one of the dashed lines to create a rectangularly shaped object. Turn the object so that the long side is vertical. Then make two triangular folds from the top left and bottom left corners along the dotted lines to produce a small funnel-shaped object. If this is dropped it will quickly settle into a nice steady fall at a terminal velocity with the point down. This is known as a relative equilibrium in that all the state variables are constant, except the vertical position, which is decreasing linearly. It is locally stable since small perturbations keep the trajectories close, and is also globally attracting in the sense that all initial conditions eventually lead to this steady falling.

If the folds are then smoothed out by flattening the paper more back to its prefolded shape, then only when the paper is dropped very carefully will it fail to flutter. This nearly flat paper has a relative equilibrium consisting of flat steady falling, but the basin of attraction of this equilibrium is very small. That is, the more folded the paper is, the larger the set of initial conditions that will lead to steady falling. If the folds are sharp enough and the distance to the floor great enough, then no matter how the paper is dropped it will eventually orient itself so the point is down, and then fall steadily.

This large change in qualitative behavior as a parameter of the system is changed (in this case, the degree of folding) is the subject of bifurcation analysis

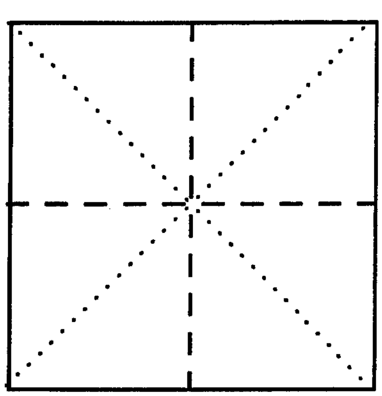

FIGURE B.2 Proper folding diagram for bifurcation experiment.

within the theory of dynamical systems theory. In these examples, bifurcation analysis could be used to explore why a regular coin shows STIC only after the first collision, while the paper coin shows it only up to hitting the floor, as well as why the dynamics of the folded paper change with the degree of folding. Of course, bifurcation analysis applies to mathematical models, and developing such models for these examples is not trivial. To develop models that reproduce the qualitative behavior we see in these simple experiments requires advanced undergraduate level aerodynamics. These models will necessarily be nonlinear if they are to reproduce the fluttering motion, as this requires a nontrivial nonlinear model for the fluids.

Bifurcation is related to chaos in that bifurcation analysis has often been an effective tool to study how complex systems transition from regular behavior to chaotic behavior. While chaos per se may be overrated, the underlying concepts of sensitivity to initial conditions and bifurcation, and more generally the role of nonlinear phenomena, are critical to the understanding of complex systems. The bottom line is as follows:

- We can make models from components that are simple, predictable, deterministic, symmetric, and homogeneous, and yet produce behavior that is complex, unpredictable, chaotic, asymmetric, and heterogeneous.
- Of course, in engineering design we want to take components that may be complex, unpredictable, chaotic, asymmetric, and heterogeneous and interconnect them to produce simple, reliable, predictable behavior.

We believe that the deeper ideas of dynamical systems will be important ingredients in this effort.

Complexity

It is tempting to view complexity in this context as something that arises in a mystical way between complete order (that the coin will come to rest) and complete randomness (heads or tails) and to settle on chaotic systems as prototypically complex. We prefer to view complexity in a different way. To make reliable predictions about, say, the final horizontal resting place, the distribution of horizontal resting positions, or the distribution of trajectories, we would need elaborate models about the experiment and measurements of properties of the mechanism, the coin, and the floor. We might also improve our prediction of, say, the horizontal resting location if we had a measurement of the positions and velocities of the coin at some instant after being tossed. This is because our suspicion would be that the greatest source of uncertainty is due to the tossing mechanism, and the uncertainty created by the air and the collisions with the floor will be less critical, but this would also have to be checked. The quality of the measurement would obviously greatly affect the quality of any resulting prediction, of course.

To produce a model that reliably predicted, say, the distribution of the trajectories could be an enormous undertaking, even for such a simple experiment. We would need to figure out the distributions of initial conditions imparted on the coin by the tossing mechanism, the dynamics of the trajectories of the coin in flight, and the dynamics of the collisions. The dynamics of the coin in the air is linear if the fluid/coin interaction is ignored or if a naive model of the fluid is assumed. If the coin is light, and perhaps flexible, then such assumptions may allow for too much uncertainty, and a nonlinear model with dynamics of the coin/fluid interaction may be necessary (imagine a "coin" made from thin paper, or replace air by water as the fluid). If the coin flexibility interacts with the fluid sufficiently, we could quickly challenge the state of the art in computational fluid dynamics.

The collisions with the floor are also tricky, as they involve not only the elastic properties of the coin and floor, but the friction as well. This now takes us into the domain of friction modeling, and we could again soon be challenging the state of the art. Even for this simple experiment, if we want to describe detailed behavior we end up with nonlinear models with complex dynamics and the physics of the underlying phenomena is studied in separate domains. It will be difficult to connect the models of the various phenomena, such as fluid/coin interaction, and the interacting of elasticity of the floor and coin with frictional forces. It is the latter feature that we refer to as heterogeneity. Heterogeneity is mild in this example since the system is purely mechanical, and the collisions with the floor are relatively simple.

Our view of complexity, then, is that it arises as a direct consequence of the introduction of dynamics, nonlinearities, heterogeneity, and interconnectedness intended to reduce the uncertainty in our models so that reliable predictions can be made about some specific behavior of our system (or its unpredictability can be reliably confirmed in some specific sense, which amounts to the same thing). Complexity is not an intrinsic property of the system, or even of the question we are asking, but in addition is a function of the models we choose to use. We can see this in the coin tossing example, but a more thorough understanding of complexity will require the richer examples studied in the rest of this paper.

While this view of complexity has the seemingly unappealing feature of being entirely in the eye of the beholder, we believe this to be unavoidable and indeed desirable: Complexity cannot be separated from our viewpoint.

Uncertainty Modeling and Management

Up to this point, we have been rather vague about just what is meant by uncertainty, predictability, and complexity, but we can now give some more details. For our coin toss experiment, we would expect that repeated tosses would produce rather different trajectories, even when we set up the tossing mechanism identically each time to the extent we can measure. There would

presumably be factors beyond our control and beyond our measurement capability. Thus any model of the system that used only the knowledge available to us from what we could measure would be intrinsically limited in its ability to predict the exact trajectory by the inherent nonrepeatability of the experiment. The best we could hope for in a model would be to reliably predict the possible trajectories in some way, either as a set of possible trajectories or in terms of some probability distribution. Thus we ideally would like to explicitly represent this uncertainty in our model. Note that the *uncertainty is in our model* (and the data that goes with it). It is we—not nature—who are uncertain about each trajectory.[1] We now describe informally the mechanisms by which we would introduce uncertainty into our models.

Parametric Uncertainty

A special and important form of uncertainty is parametric uncertainty, which arises in even the simplest models such as attempting to predict the detailed trajectory of a coin. Here the "parameters" include the coin's initial conditions and moments of inertia, and the floor's elasticity and friction. Parameters are associated with mechanisms that are modeled in detail but have highly structured uncertainty. Roughly speaking, *all* of the "inputs" to a simulation model are parameters in the sense we use the term here.[2]

How do we deal with parametric uncertainty (see also Appendix D)?

- *Average case.* If only average or typical behavior is of interest, this can be easily evaluated with a modest number of repeated Monte Carlo simulations with random initial conditions. In this case the presence of parametric uncertainty adds little difficulty beyond the cost of a single simulation. Also, in the average case the number of parameters does not make much difference, as estimates of probability distributions of outcomes do not depend on the number of parameters.
- *Linear models.* If the parameters enter linearly in the model, the resulting uncertainty is often easy to analyze. To be sure, we can have extreme sensitivity to initial conditions, but the consequences are easily understood. Consider the linear dependence of the velocity and position of the first floor collision as a function of the initial velocities and positions of the coin. A set in the initial

[1]Except in our discussion of VLSI later in this appendix, we ignore quantum mechanics and the intrinsically probabilistic behaviors associated with it. Quantum effects are only very rarely significant for the systems of interest here.

[2]Some workers distinguish between "parameters" that can be changed interactively at run time, or in the course of a run, and "fixed data," that can be changed only by recompiling the database. Both are parameters for the purposes of this paper.

condition space is easily mapped to a set of collision conditions. Both average and worse-case behavior can be evaluated analytically. For example, suppose we have some scalar function $F(p)$, where p is some vector of n parameters. If F is linear, then evaluating its largest values over a convex set of parameters is straightforward, and in the case where each component of the parameters is in an interval, it is trivial.

- *Nonlinear/worst-case/rare event.* If the parameters enter nonlinearly, then the analysis becomes more difficult—particularly if we are interested in worst-case behavior or rare events. For example, if $F(p)$ is nonlinear and has no exploitable convexity properties, we may have to do a global and exhaustive search of p to determine, say, the maximum value of F. Such a search will grow exponentially with the dimension n of p. While this is an entirely different role of exponential growth than in sensitivity to initial conditions, the consequences are no less dramatic. If we only choose to examine a gridding of the space of parameters where we take 10 values of each element (this could be too coarse to find the worse case), then the number of functional evaluations will be 10^n. No matter how quickly we do the functional evaluations, this exponential growth prohibits exploring more than a handful of parameters.

It does not take highly nonlinear dynamics to produce nonlinear dependence on parameters. A simplified model of coin flipping with a linear model for collisions and a linear model for flight between collisions is piecewise linear, but this is enough to produce complicated dependence on initial conditions. In a model with linear dynamics, the dependence of final conditions on initial conditions is linear, but parameters such as mass, moments of inertia, resistance, and so on, can enter the equations nonlinearly, thereby producing a nonlinear dependence of the trajectories on these parameters.

Evaluating worst-case or rare events can also be more difficult than the average case, because Monte Carlo requires an excessive number of trials or we have to exhaustively search all parameters for the worst case. This too can be computationally intractable, depending on the number of the parameters and the functional dependence on them. In some cases, exact calculation of probability density functions can be easier, and numerical methods for evaluating probability density functions by advanced global optimization methods is a current research topic.

Of course, we would always hope to discover some property of the parametric dependence that makes it possible to avoid this exponential growth in computation with the number of parameters. Unfortunately, it has been proved that evaluating, say, the stability of even the simplest possible nontrivial linear system depending on parameters is NP-hard, which implies that exponential growth is in a certain sense unavoidable. To explain the implications of this for evaluating parametric uncertainty will require several additional concepts, including obviously the meaning of NP-hard, and will be taken up in more detail later.

Noise and Unmodeled Dynamics

Not all uncertainty is naturally modeled as parametric. Suppose the fluid in which our coin is tossed cannot be neglected, as would be the case for a light paper coin in air, or a coin in water. The coin may exhibit erratic and apparently random motion as it falls. While we could in principle model the fluid dynamics of the atmosphere in detail, the error-complexity tradeoff is very unfavorable. Traditionally, noise is instead modeled as a stochastic process even when it is more appropriate to think of it as chaotic. That is easy to see in the case of wind gusts generated by fluid motions in the air. Just because the wind appears random and we do not want to model the fluid in detail does not mean that wind is naturally viewed as a random process. We model it as a stochastic process for convenience.

Other examples of phenomena that are often modeled as noise include arrival times for queues, thermal noise in circuits, and even the outcome of the process of coin flipping itself. In each case, the mechanism generating the noise is typically not modeled in detail, but some aggregate probabilistic or statistical model is used. Robust control has emphasized set descriptions for noise, in terms of statistics on the signals such as energy, autocorrelations, and sample spectra, without assuming an underlying probability distribution. Recently, the stochastic and robust viewpoints have been reconciled by the development of set descriptions that recover many of the characteristics and much of the convenience of stochastic models. Finally, Monte Carlo simulations adequately generate noise with pseudo-random number generators, which are neither truly chaotic nor random, but are periodic with very long periods.[3] Thus, it makes little sense to be dogmatic about insisting that models be stochastic, per se, but it is important that noise sources be explicitly modeled in some reasonable way. This may be accomplished in a number of ways:

- *Unmodeled dynamics.* The use of constant parameters or noise to model aerodynamic forces generated by the fluid around our coin means not treating explicitly the complex, unsteady fluid flows that more accurately describe the physics. Even if we attempt to model the fluids in some detail, the forces and moments that the resulting model predicts will be "felt" by the coin will be wrong, and the difference may depend on the coin's movements. It may be undesirable to then model these forces as noise, since that assumes that they do not depend on the coin's movement. We may choose to model this type of uncertainty with a mechanism similar to noise above, but involving relationships between variables, such as coin velocities and forces. Like noise modeling, we

[3]Some MC simulations have been done with genuinely random physical sources, such as decay of radioactive isotopes.

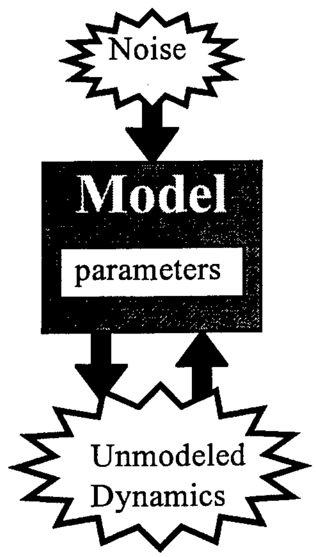

FIGURE B.3 Relationships of models and their parameters with noise and unmodeled dynamics. Noise excites model dynamics but not vice versa.

use bounds and constraints, but now on the signals describing both the coin and the fluid. Similarly, rigid body models assume that forces directly generate rigid body motion, while models that included flexible effects allow for bending as well. Rather than modeling the flexibility in detail, we may choose to bound the error between forces and rigid body motions as constraints on signals.[4]

Noise, then, is similarly a special case of unmodeled dynamics where one assumes that the unmodeled dynamics excite the modeled dynamics but not vice versa (see Figure B.3). For example, modeling atmospheric gusts as noise assumes that the vehicle or coin motion has a negligible effect on the atmosphere compared to the fluid motions in the atmosphere generated by other sources. This is a reasonable assumption in many circumstances but might be unreasonable for, say, airplanes flying near large fixed objects such as the ground.

[4]This issue of how to treat unmodeled dynamics comes up in quasi-steady aerodynamics of airplanes, which is a particular problem when considering the interactions of closely spaced vehicles whose flows affect each other, or when the vehicle is highly flexible. It is currently beyond the state of the art to do computational fluid dynamics (CFD) for a moving vehicle in a way that would allow the use of CFD codes in vehicle simulation. It is possible to put explicit uncertainty models into the coefficients and analyze their effects using methods from robust control. In fact, exactly this type of analysis for the Shuttle Orbiter during reentry and landing was among the first commercial successes of robust control, and the use of robustness analysis software is now standard in the Shuttle program. CFD, robust control, and the Shuttle reentry will be discussed in more detail later.

• *Sensitivity to uncertainty.* While we have discussed STIC with simple examples, models can be sensitive to all forms of uncertainty, not just initial conditions. This is a much more subtle notion and requires, unfortunately, deeper mathematics to explain. Indeed, as we will see from the engineering example later, sensitivity to unmodeled dynamics is a much greater problem generally than sensitivity to initial conditions.

To get some sense of the issues, here, suppose that we develop a model of the paper falling that assumes the paper is rigid and the flow is very simple, and that it seems to capture the qualitative dynamics of the experiment. Suppose we then try to use this model to predict, *before* doing any experiment, what would happen with tissue paper, where there is substantial flexibility. The behavior might be totally different, and no choices of parameters in our simpler model would capture the behavior of the falling tissue. One way to approach the tissue would be to try initially to bound the effects of flexibility in a rough way and check if this makes any difference to the outcome (assume for the moment we have some tools to do this). Presumably, this would reveal that the small flexibility makes little difference for the bifurcation analysis with regular paper, but makes a large difference with the tissue. The uncertain model of the tissue would suggest that we would be unable to reliably predict details of the trajectory as well as in the regular paper case without doing a more detailed model of the tissue flexibility. Thus, our initial model of the tissue would have been sensitive to unmodeled dynamics.

• *Games with hostile adversaries.* This appendix deals largely with "hard" engineering examples, but before leaving the subject of uncertainty, it is worth noting a complication of particular importance in military applications—notably, the presence of a hostile adversary. It is a complication without direct parallel in normal science and engineering. Nature, while perhaps sometimes seeming to be capricious, does not consciously plan how to complicate our lives. In conflict, all the participants may have strategies that change in response to perceived actions of the other participants. These changes may seek to optimize some feature of events, or may merely be dictated by doctrine. They may be objectively "optimal" in some sense, or they may be idiosyncratic and risk-taking. Some adversaries may wish to optimize the likelihood of complete success, caring little about "expected value." And all humans are subject to well-known cognitive biases that influence decisions.

Now, some of the methods discussed in this appendix might well be useful in modeling potential adversary behaviors. We could, for example, model all participants' options and reasoning, and then look at, say, minimax strategies. Indeed, robust control theory often models parameters, noise, and unmodeled dynamics in such a way that control design can be viewed as a differential game between the controller and "nature." Such methods could be quite useful for designing robust strategies. We will not pursue this subject further here, except

to notice parenthetically that it is much easier to produce a computer program that plays grandmaster-level chess (which has been done) than it is to model accurately the actual play of a grandmaster (which might not be done in the foreseeable future).[5]

Uncertainty Management

The motivation for introducing uncertainty mechanisms into our models is to predict the range and distribution of possible outcomes of our experiments. What is somewhat misleading about coin tossing with respect to the broader VE area is the small scale of the experiment and its relative homogeneity. In VE modeling we must expect huge numbers of components with extremely diverse origins. We have discussed the features of a model that included nonlinearity, dynamics, interconnection of heterogeneous components, complexity, and uncertainty. Of these, only uncertainty is an intrinsic property of the modeling process, with the others introduced—perhaps reluctantly—to reduce the uncertainty in our models. This uncertainty management is the driving force behind complexity in models.

Conventional modeling in science and engineering is basically reductionist. Experiments are designed to be controlled and repeatable, usually with certain phenomena isolated. It is widely thought that if we model a system with sufficient accuracy, then we can reliably predict the behavior of that system. This standard modeling paradigm is poorly suited to VE. The phenomenon of deterministic chaos has shed some doubt on the conventional view, and it is now widely appreciated, even among the lay public, that quite simple models can produce apparently complex and unpredictable behavior. What is less well appreciated is an entirely different issue of how uncertainty in components interacts with the process of interconnection of components to produce uncertainty at the system level: There is simply no way of telling how accurately a component must be modeled without knowing how it is to be connected. Thus both the component uncertainty and the interconnection determine the impact of that uncertainty.

One reason that explicit uncertainty modeling is uncommon is that the typical "consumer" of a conventional model is another human sharing domain-specific expertise and an understanding of standard assumptions. *Model uncertainty and the interpretation of assumptions are typically implicit and also part of the domain-specific expertise.* Most scientific and engineering disciplines can almost be defined in terms of what they choose to both focus on and neglect about

[5]Within the military domain, there are examples to illustrate all of these points. For example, RAND's SAGE algorithm, developed by Richard Hillestad, is used in theater-level models to "optimally" allocate and apportion air forces across sectors and missions. This not only establishes bounds on performance, but also reduces the number of variables, which makes it much easier to focus on tradeoffs among weapon systems and forces. This said, it is important to run cases in which the sides follow doctrine, because that behavior may be quite different and more realistic.

the world. Within domains, the dominant assumptions and viewpoints are taken for granted and never mentioned explicitly. In the coin flipping experiment, terms such as inviscid flow might be mentioned, but not something like the assumption of chemical equilibrium.

Problems with conventional modeling arise in VE because the consumer of a model will be a larger model in a hierarchical, heterogeneous model, with possibly no intervention by human experts. We cannot rely on theory or software that is domain-specific, implicit, and requires human interpretation at every level. The more complex the system, the less critical individual components may become but the more critical the overall system design becomes. Furthermore, we are interested in modeling real-world behavior, not an idealized laboratory experiment.

Model Fidelity Versus Cost

The next question we address is, What constitutes a sensible notion of model error or fidelity? Again starting with a simple "naive" view, let us imagine that reality has some set of behaviors, that our model has some set of behaviors, and that we have some measure of the mismatch between these two sets. These may be heroic assumptions, but let us press forward for now.

Putting aside the cognitive preference for simplicity, we would perhaps obviously prefer high-fidelity models. The obstacles to model fidelity are the costs of modeling due to limited resources, mainly:

- Computation,
- Measurement, and
- Understanding.

These limitations are ultimately connected to time and money. This suggests tradeoffs.

For the time being, we can think of computation cost naively as simply the cost of the computer resources to simulate our model, and measurement cost as the cost to do experiments, take data, and process those data to determine parameters in our model. It is obviously not so easy to formalize what we mean by understanding, but it is clear that we can have vastly different levels of understanding about different physical processes and this can greatly affect our ability to effectively model them.

Figure B.4 illustrates the tradeoff between fidelity and complexity, or more precisely the tradeoff between model error and resources. As we use more resources, we can reduce our error, but there are strong effects of diminishing returns.[6]

[6]Actually, errors could *increase* with cost if resources are used to build extra complexity to the point that the model begins to collapse under its own weight, but here we assume optimal use of the resources.

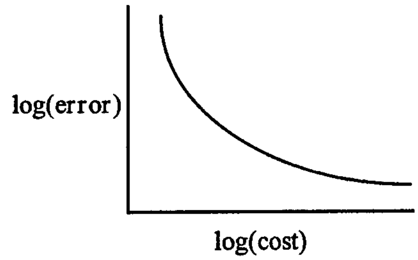

FIGURE B.4 Error-cost tradeoff.

A well-known example occurs in weather prediction. For standard computer simulations used in weather prediction, the error between the prediction and the actual weather grows with time due to sensitivity to initial conditions. Long-term predictability is viewed by experts as being impossible, and even massive improvements in computation and measurement will yield diminishing returns.

Uncertainty and Complexity Revisited

We can summarize the previous discussion as follows. No matter how sophisticated our models, there is always some difference between the model and the real world. Model uncertainty leads to unpredictability, which mirrors the unpredictability of real systems. This has two important aspects. One is that models can exhibit extreme sensitivity to variations in model assumptions, parameters, and initial conditions. The second, discussed below, is the combinatorial complexity of evaluating all the model combinations that arise from possible variations in assumptions, parameters, and initial conditions in all the subsystems, which makes a brute force enumeration prohibitively expensive. These are some of the fundamental limitations on the predictability of models, which will not be eliminated by advances in computational power, measurement technology, or scientific understanding. Thus in developing robust VE, there are certain "hard" limits on predictability, and it is important to understand and quantify the limits on the predictability of full system models in terms of the uncertainties in component models. Current VE enterprises generally do not have good strategies for dealing with these uncertainties, or for understanding how they propagate through the system model and ultimately affect the decision-making process they were intended to serve.

To make reliable predictions about the systems being modeled under uncertainty, we are often forced to add *complexity in the form of nonlinearities, dynamics, and interconnections of heterogeneous components. Unlike uncertainty, none of these are intrinsic properties of our systems and models*, but are added, perhaps reluctantly, to reduce uncertainty. It is tempting to say that all real systems are nonlinear, dynamic, and heterogeneous, but this is meaningless since

these are properties of models, not reality (although we will use the convenient shorthand of referring to some phenomena as, say, nonlinear when it might be more precise to say that high-fidelity models of the phenomena must be nonlinear). It would be more appropriate to say that modeling of physical systems leads naturally and inevitably to the introduction of such mathematical features. Complexity is due, in this view, simply to the presence of these features, and to the degree of their presence. This view of complexity will be a theme throughout, so let us consider a bit more detail on the meaning of dynamics, nonlinearity, and heterogeneity.

From Static to Dynamic Systems

A dynamic model is simply one whose variables evolve with time, perhaps described by differential or difference equations, or perhaps more abstractly. For partial differential equations, the situation is even more complicated, since solutions can vary continuously with space as well as time. Thus, dynamical systems tend to be much more difficult to work with than static systems.

From Linear to Nonlinear Systems

The importance of linear models is that they satisfy superposition, so that a linear function of a list of variables can be completely characterized by evaluating the function on one value of each variable taken one at time. In other words, the local behavior of a linear function completely determines its global behavior. Nonlinearity is the absence of this property, so that local information may say nothing about global behavior. What is critical in modeling uncertain systems is how the quantity that we want to predict depends on the uncertainty we are modeling. A linear dynamical system can depend nonlinearly on parameters in such a way as to make evaluation of the possible responses of the system very difficult. In turn, a nonlinear dynamical system may have outputs that depend only linearly on some parameters, and this may be easily evaluated. What is critical is the dependence on the uncertainty.

Again, nonlinearities are not responsible for sensitivity to initial conditions. It is when nonlinearities are combined with sensitivity to initial conditions that behavior can be a complex and unpredictable function of the initial conditions.

From Homogeneous to Heterogeneous Systems

Heterogeneity in modeling arises from the presence in complex systems of component models dealing with material properties, structural dynamics, fluids, chemical reactions, electromagnetics, electronics, embedded computer systems, and so on, each represented by a separate and distinct engineering discipline. Often, modeling methods in one domain are incompatible with those in another.

Even when we can break the system into multiple levels with the lowest level containing only one modeling domain, the component models must be combined. Thus, a critical issue that must be dealt with in every aspect of modeling is how to combine heterogeneous component models.

A classical example of a mildly heterogeneous system occurs in the phenomenon of flutter, an instability created by the interaction of a fluid and an elastic solid, or more generally with a solid that has nontrivial dynamics. It is a critical limiting factor in the performance of many aircraft wings and jet engines, and a simple version of it could be seen in the "flutter" of the paper in the experiment above. There are two approaches to treating such heterogeneous systems. One is to do a fully integrated model of the system, which in this case is the domain of the field of aeroelasticity. The other is to make simplifying approximations about the boundary conditions between the heterogeneous components and then interconnect them. For flutter, this is typically done when assuming that the fluid is treated quasi-statistically, so forces and moments generated by the fluids do not depend on the dynamics of the solid material. This way, the *dynamics* of the fluid can be treated separately, approximated as a static map, and then simply connected with the dynamics of the solid. If this approximation is reasonable, as it is in our paper experiment, then simple models can predict the presence or absence of flutter.

Flutter is only a mildly heterogeneous system, because it involves continuum mechanical phenomena, but, for example, no chemistry, electromagnetics, or thermodynamics. In more profoundly heterogeneous systems, it is not possible to create huge new engineering domains to address the unique modeling problems associated with each possible combination of modeling problems. We must make suitable approximations of the boundary conditions between domains/components.

From Isolated Systems to Systems That Are Interconnected, Hierarchical, Multi-resolution, Variable Granularity, and Variable Fidelity

In simple situations it is often possible to use aggregated models that capture the system behavior without detailed treatment of the system's internal mechanisms. However, in developing detailed models of complex systems, it is common practice to break the system up into components, which are then modeled in detail. This can continue recursively at several levels to produce a hierarchical model, and the full system must then be built up from the component models.

This disaggregation approach is essentially the only way that complex system models can be developed, but it leads to a number of difficult problems, including the need to connect heterogeneous component models, and the need for multi-resolution or variable granularity models. Perhaps more importantly, this neo-reductionist approach to modeling may represent a reasonable scientific program for discovering fundamental mechanisms, provided one never wants to

reconstruct a model of the whole system. Applied naively, it simply does not work very well as a strategy for modeling complex systems.

There is need for multi-resolution or variable granularity models, because the process of component disaggregation can continue indefinitely to create an infinitely complex model (see also Appendix E). It is possible, and even likely, that this process will not converge, as there can easily be extreme sensitivity to small component errors that are the consequence of interconnection and that can be evaluated only at the system level. For example, there might exist quite simple models for components and their uncertainty that interconnect to reliably predict system behavior, but it can be essentially impossible to discover these models by viewing the components in isolation. Thus one is in the paradoxical position of needing the full system to know what the component models should be, but having only the component models themselves as the route to creating a system model. Neither a purely top-down nor a bottom-up approach will suffice, and more subtle and iterative approaches are needed. Even the notion of what is a "component" is not necessarily clear a priori, as the decomposition is never unique and often the obvious thing to do is severely suboptimal.

The standard approach to developing variable error-complexity component models is to allow multi-resolution or variable granularity models. Simple examples of this include using adaptive meshes in finite element approximations of continuum phenomena, or multi-resolution wavelet representations for geometric objects in computer graphics. In hierarchical models, the problem of developing analogous variable resolution component models is quite subtle and will surely be an important research area for some time. Using these same examples, there are difficult problems associated with modeling continuum phenomena involving fluid and flexible structure interaction, as well as phase changes. Similarly, building aggregate multi-resolution models of interconnected geometric objects from multi-resolution component models is a current topic of research in computer graphics.

CASE STUDIES IN SUCCESSFUL VIRTUAL ENGINEERING

With this background of basic concepts, let us now consider some case studies to illustrate successful VE. Advocates of the power of modeling and simulation typically extrapolate from three shining examples, which could be thought of as "proto-VE" case studies: the computer-aided design (CAD) of the Boeing 777, computational fluid dynamics (CFD), and very large scale integrated circuits (VLSI). Each has made great use of computation and has challenged in various ways available hardware and software infrastructures. While these success stories are certainly encouraging, great caution should be used in extrapolating to more general situations, as a review of the state of the art will reveal. Indeed, each of these domains has serious limitations and faces major challenges. Examples of major failures of complex engineering systems (e.g., the Titanic, the

FIGURE B.5 The Boeing 777. SOURCE: Boeing Web site (www.boeing.com/companyoffices/gallery/images/commercial/).

Tacoma Narrows bridge, the Denver baggage handling system, and the Ariane booster) should cause us all to be sobered. We will argue that *uncertainty management* together with *dynamics and interconnection* are the key to understanding both these successes and failures and the future challenges.

Boeing 777

Boeing invested more than $1 billion (and insiders say much more) in CAD infrastructure for the design of the Boeing 777 (see Figure B.5), which is said to have been "100 digitally designed using three-dimensional solids technology." Boeing based its CAD system on CATIA (short for Computer-aided Three-dimensional Interactive Application) and ELFINI (Finite Element Analysis System), both developed by Dassault Systemes of France and licensed in the United States through IBM. Designers also used EPIC (Electronic Preassembly Integration on CATIA) and other digital preassembly applications developed by Boeing. Much of the same technology was used on the B-2 program.

While marketing hype has exaggerated aspects of the story, the reality nonetheless is that Boeing reaped huge benefits from design automation. The more

than 3 million parts were represented in an integrated database that allowed designers to do a complete 3D virtual mock-up of the vehicle. They could investigate assembly interfaces and maintainability using spatial visualizations of the aircraft components to develop integrated parts lists and detailed manufacturing process and layouts to support final assembly. The consequences were dramatic. In comparing with extrapolations from earlier aircraft designs such as those for the 757 and 767, Boeing achieved the following:

- Elimination of >3,000 assembly interfaces, without any physical prototyping,
- 90 percent reduction in engineering change requests (6,000 to 600),
- 50 percent reduction in cycle time for engineering-change request,
- 90 percent reduction in material rework, and
- 50-fold improvement in assembly tolerances for the fuselage.

While the Boeing 777 experience is exciting for the VE enterprise, we should recognize just how limited the existing CAD tools are. They deal only with static solid modeling and static interconnection, and not—or at least not systematically—with dynamics, nonlinearities, or heterogeneity. The virtual parts in the CATIA system are simply three-dimensional solids with no dynamics and none of the dynamic attributes of the physical parts. For example, all the electronics and hydraulics had to be separately simulated, and while these too benefited from CAD tools, they were not integrated with the three-dimensional solid modeling tools. A complete working physical prototype of the internal dynamics of the vehicle was still constructed, a so-called "iron-bird" including essentially everything in the full 777.

While there was finite element modeling of static stresses and loads, all dynamical modeling of actual flight, including aerodynamics and structures, was done with "conventional" CFD and flight simulation, again with essentially no connection to the three-dimensional solid modeling. Thus while each of these separate modeling efforts benefited from the separate CAD tools available in their specialized domains, this is far from the highly integrated VE environment that is envisioned for the future, and is indeed far from even some of the popular images of the current practice. Thus while a deeper understanding of the 777 does nothing to reduce our respect for the enormous achievements in advancing VE technology or dampen enthusiasm for the trends the 777 represents, it does make clear the even greater challenges that lie ahead.

What are the next steps in CAD for projects like the 777? Broadly speaking, they involve much higher levels of integration of the design process, both laterally across the various subsystems, and longitudinally from preliminary design, through testing, manufacturing, and maintenance. They will require more systematic and sophisticated treatment of uncertainties, especially when dynamics

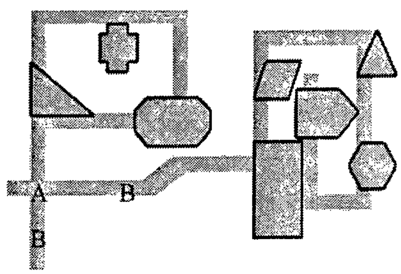

FIGURE B.6 Two multicomponent subassemblies in the 777 connecting only at point A.

are considered in a unified way. This will require introducing nonlinearities, heterogeneity, and variable resolution models.

Boeing engineers view these steps as enormous challenges that must be faced. Even something as simple-sounding as using the CATIA database describing the three-dimensional layout of the hydraulics and their interconnections as a basis for a *dynamic* simulation of the hydraulics remains an open *research* problem, let alone using CATIA as a basis for dynamic modeling and simulation of aerodynamics and structures. What is difficult to appreciate is how the sheer scale of keeping track of millions of components can be computationally and conceptually overwhelming.

Interference Analysis

To illustrate some of the issues in the three-dimensional solid modeling for the 777, consider yet another simple experiment in two dimensions. Suppose that we have two two-dimensional subassemblies, each consisting of several components, that we wish to interconnect at point A as shown in Figure B.6 (the components shown have no meaning and are simply for illustration.) We want to be sure there are no unwanted intersections in the design, and it is clear from Figure B.6 that this assembly has no connections except at point A.

The 777 has millions of such parts. A virtual mock-up can be made from a parts and interconnection list so that designers can "fly through" the design to check for unwanted interconnections. The computer can also automatically check for such interferences so that these can be identified and redesigned before they are discovered (more dramatically and at much greater expense) during physical assembly. If there are n components, we can think of an $n \times n$ matrix of pairs of potential collisions, so 3,000,000 parts would have approximately $n*(n-1)/2 = 4.5 \times 10^{12}$ possible intersections to be checked. Although this grows only quadratically with the number of parts (not the exponential growth we are so concerned with elsewhere), the sheer number of parts makes brute force enumeration unattractive. Fortunately, there are standard ways to reduce the search.

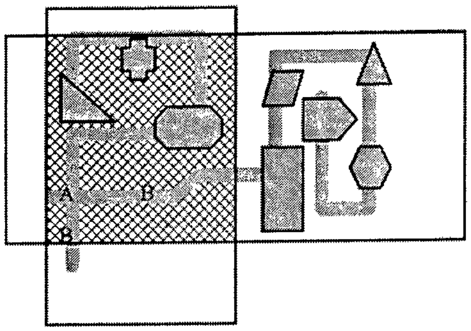

FIGURE B.7 Simple bounding boxes with large intersection and not conclusively eliminating unwanted subassembly interconnection.

We could begin by putting large bounding boxes around the subassemblies, as shown in Figure B.7. This could be used to eliminate potential intersections far away from these subassemblies (resulting in large sections of our interconnection matrix that would not need to be checked), but would not conclusively eliminate unwanted connections between the subassemblies. At this point, all the pairwise components of the subassemblies could be checked, or we could refine the bounding boxes, as shown in Figure B.8. At this point, we would have eliminated all but 2 components, and they could be checked to see that the only intersection was at A. The bounding boxes in this case reduced the cost from computing 24 pairwise (4×6) intersections to computing 1 pairwise component and a few bounding boxes. The bounding boxes have simple geometries, so are more easily checked than the components, but need to be constructed. Clearly, there is a tradeoff, and one does not want to use too many or too few bounding boxes.

This is an example of a general technique for searching called divide and conquer, where the problem is broken up into smaller pieces using some heuris-

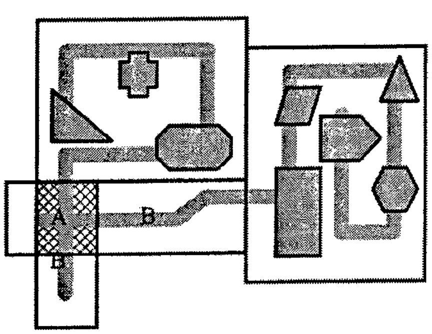

FIGURE B.8 Refined boxes show only intersection is at A.

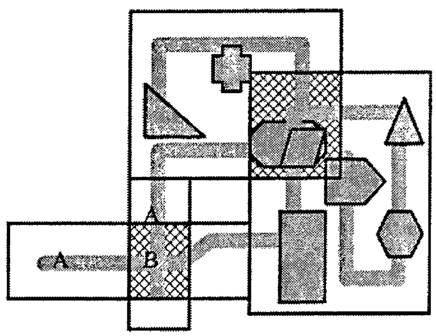

FIGURE B.9 Refined boxes help find interferences.

tic. It is related to branch and bound, where, say, a function to be minimized is searched by successively breaking its domain into smaller pieces on which bounds of the function are computed. Suppose we want to compute the minimum distance between components that are not supposed to be connected (we want to make sure this function is bounded away from zero). A bounding box gives us upper and lower bounds on this function for the component combinations that are included in the bounding boxes. We can ignore pairs of boxes that are separated by more than the smallest upper bound we have, thus pruning the resulting tree of refined bounding boxes. This is illustrated in Figure B.9, where the subassemblies are connected at point B, instead of A. The refined bounding boxes show how they can be used to focus the search for unwanted interconnections.

Note that if we introduce uncertainty in our description of the components, it can drastically increase the computation required to do pairwise checking and make the bounding box approach even more attractive. Actually, while the basics of solid modeling are well developed, there is no standard approach to uncertainty modeling even here and many open questions. Once we introduce uncertainty in a general way, then exponential growth in evaluating all the possibilities becomes a worry. Because three-dimensional solids inherently have limited dimensionality of contacts, it should be possible to avoid this. As we will see later, uncertainty in dynamical systems is even more challenging, and a version of the bounding box idea is quite useful in doing robustness analysis of uncertain *dynamical* systems as well.

Dynamic Simulations of Aircraft Flight

It is interesting to note that the Boeing 777 used an advanced, though conventional, approach to modeling and simulation of the aerodynamics and flight. Aeromodeling is a convenient example because there is a long history of successful systematic modeling, yet substantial challenges remain. It also offers us a

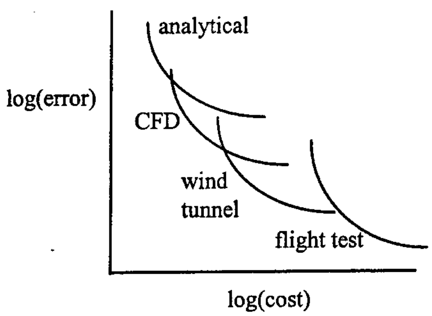

FIGURE B.10a Schematic of error-cost relationship of different approaches to obtaining aerocoefficients for flight modeling.

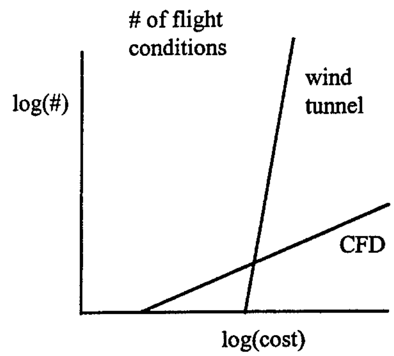

FIGURE B.10b Cost versus number of points at which aerocoefficients are obtained.

chance to discuss computational fluid dynamics (CFD) in a broader context that includes analytic tools, wind tunnels, and flight test. This is simply for illustration, but it will touch on broader issues in VE.

Standard rigid body airplane models typically consist of a generic form of the model as an ordinary differential equation (ODE) with vehicle-specific parameters for mass distributions, atmospheric conditions (dynamic pressure), and aerodynamic coefficients. This rigid body model determines what motion of the vehicle will result from applied forces due to propulsion and aerodynamics. Aerocoefficients are parameters that give the ratio of forces and moments generated on the vehicle to surface deflections and angular changes of the vehicle with respect to the ambient air flow. The standard quasi-steady assumption is that these aerocoefficients depend only on sideslip angle and angle of attack, and not on the history of the motion of the vehicle or the complex flow around it. Therefore the aerocoefficients are most compactly represented as a set of six functions, three forces and three moments, on the unit sphere (the two angles), plus additional functions for surface deflections. Obtaining these functions dominates standard aeromodeling and the use of CFD.

Aerocoefficients can be estimated analytically, computed using CFD, or measured in wind tunnels or flight tests. Figure B.10a shows schematically the error-costs associated with the various methods. Once the coefficients are obtained, they can be plugged into a dynamic model of the airplane, and the resulting nonlinear differential equations can be simulated. This figure is misleading in many ways, one of which is that the error-cost tradeoff between CFD and wind tunnel is oversimplified. Figure B.10b shows the cost to find the aerocoefficients at a certain number of points. Building a wind tunnel model is relatively expensive, but once it is available, the incremental time and cost to do an additional experiment are small. Research is being done to speed up both CFD and the

building of wind tunnel models, thus shifting both curves to the left. One of the most revolutionary developments currently going on in this area is the improvement in rapidly generating wind tunnel models from computer models. Researchers are currently working to change the traditional turnaround time from months to hours.

These figures do not address the fact that CFD and wind tunnels do not give exactly equivalent results. To a first approximation, wind tunnel tests can give lower overall modeling error by allowing the modeler to include additional factors, such as unsteady effects. On the other hand, CFD can provide detailed flow field information that is difficult to obtain experimentally and avoids experimental artifacts like wind-tunnel wall effects. Finally, flight tests give the most reliable predictions of aircraft behavior, although even here there are errors, since not all possible operational conditions can be tested or even necessarily anticipated. This is summarized in Figure B.10a, which shows the error versus complexity for various modeling methods, where, vaguely speaking, error is the difference between predicted operational behavior and actual operational behavior, and cost could be taken as the total dollar cost to achieve a given error with a specific method.

Note that the methods are complementary, each best for some particular error-cost level. They are also complementary in other ways, as the deeper nature of the errors is different as well. While the details may change with technology, the overall shape of this figure will not. One goal of VE is to reduce the error associated with simulation-based methods and thus reduce the need for wind tunnel and flight testing. If this is not done carefully, it is quite easy to simultaneously increase error and cost.

This discussion has taken a very superficial view of modeling and particularly of uncertainty, but has hopefully illustrated the tradeoff between error and cost that holds across both virtual and physical prototyping. One important point to note is that despite earlier euphoric visions of the role of CFD in aircraft design that suggested it would almost entirely replace wind tunnels, only a tiny fraction of the millions of aerodynamic simulations generated for a modern aircraft design are done using CFD. The remainder continue to be done with physical models in wind tunnels, and this is not expected to change in the foreseeable future. To get a slightly deeper picture of these issues, we need to examine CFD more closely.

Computational Fluid Dynamics in Aircraft Design

Basic Elements

Our second case history involves computational fluid dynamics (CFD). Fluid dynamics is a large and sophisticated technical discipline with both a long history of deep theoretical contributions and a more recent history of major technological

impact, so it is not surprising that it is poorly understood by outsiders. We focus on the aspects relevant to the broader VE enterprise.

Rather than following a particular group of molecules, fluid dynamical models adopt a continuum view of the fluid in terms of material elements or volume elements through which the material moves. Because of the simplicity of Newtonian fluids (those for which viscosity is constant, such as air flowing about an airplane), a fairly straightforward system of partial differential equations relate the dynamics of a fluid flow element to its local velocity, density, viscosity, and externally acting forces. These are the Navier-Stokes (N-S) equations, which have been known for more than a century and a half. They merely express conservation of mass and the other of momentum, but the three-dimensional case has 60 partial-derivative terms.

The N-S equations are thought to capture fluid phenomena well beyond the resolution of our measurement technology. Thus fluid dynamics holds a very important and extreme position in VE as an example of a domain where the resource limitation is due primarily to computation and measurement, rather than to ignorance about the phenomena (although this is not true for granular or chemically reacting flow). As one might expect, then, a major effort in fluid dynamics involves various numerical approximations to solving the N-S or related equations. Such numerical airflow simulations are the subject of computational fluid dynamics (CFD). The obvious approach is direct numerical solution (DNS) of a discrete approximation to the N-S equations. Unfortunately, turbulence makes this extremely difficult.

In effect, *turbulence* is a catch-all term for everything complicated and poorly understood in fluid flow. While there is no precise definition of turbulence, its general characteristics include unsteady and irregular flows that give something of the appearance of randomness; strong vorticity; stirring and diffusion of passive conserved quantities such as heat and solutes, and dissipation of energy by momentum exchange. Under typical aircraft flight conditions at high subsonic speeds, turbulence takes the form of a nested cascade of eddies of varying scale, ranging from on the order of meters to on the order of tens of micrometers; a span of 4 or more orders of magnitude. On average, the largest eddies take energy from the free flow and, through momentum exchange, feed it down, step by step in the cascade of eddies, to the smallest eddies, where it is dissipated as heat. However, it is also possible for energy to feed from smaller eddies to larger over limited times or regions, and these reverse energy flows can play a significant role. Turbulence is no more random than the trajectories of our coins, but its sensitivity to initial conditions is even more dramatic, since there are so many more degrees of freedom. Turbulence is considered one of the classic examples of chaotic dynamics, although this has not been proved.

To fully to capture the dynamics of the airflow in DNS, it is necessary to integrate numerically over a mesh fine enough to capture the smallest turbulent eddies but extensive enough to include the aircraft and a reasonable volume of air

about it. This is beyond current computational facilities. To overcome this, various approximations must then be made, that are at the heart of CFD. Here merely make a few observations. First, as was noted earlier, only a tiny fraction of the millions of aerodynamic simulations generated for a modern aircraft design are done using CFD. The remainder are still done with physical models in wind tunnels, and this is not expected to change soon. Second, CFD is used primarily to compute the static forces on objects that are fixed relative to the flow, and any dynamical vehicle motion combines these static forces with vehicle kinematics. Using CFD for computing dynamically the forces on objects that are themselves moving dynamically adds substantially to the computational complexity.

Finally, the various approaches to CFD result in widely varying computational requirements, yet there is no integrated "master" model other than the N-S equations themselves, and substantial domain-specific expertise is needed to create specific simulation models and interpret the results of simulations. Some approximations assume no viscosity and others large viscosity; some approximations focus on material elements and their motion (called Lagrangian formulations); others focus on volume elements through which material passes, and thus fluid velocities are the focus (called an Eulerian formulation); and still others try to track the movement of the larger-scale vortical structures in the fluid. The choices are dominated by the boundary conditions, as the fluid (air) being modeled in each case is identical, and even different approximations may be made in different parts of the same flow. For example, viscosity might be modeled only near a solid boundary, while the flow far from the boundary would be assumed to be inviscid. Thus, while the material itself is perfectly homogeneous, inhomogeneities arise in our necessary attempts to approximate the fluids.

Uncertainty Management in Commercial Aircraft Design

What is particularly interesting for this appendix is the fact that, according to Paul Rubbert,[7] the chief aerodynamicist for Boeing, uncertainty management has become the dominant theme in practical applications like aircraft design. He claims that uncertainty management is replacing the old concept of "CFD validation." He argues that both CFD and wind tunnels are "notorious liars," yet modern aerodynamic designs are increasingly sensitive to small changes and error sources. Thus, better attention needs to be paid to modeling uncertainty and its consequences. CFD and wind tunnels are complementary, and the goal is to be able to reconcile the differences between their results and not to expect that they should give the same results. In this context, CFD users must be provided with the insight and understanding to allow them to manage the various sources

[7]Private communication with John Doyle, California Institute of Technology, December 1996.

of uncertainty that are present in their codes, and to understand how those uncertainties affect the specific aircraft behavior they are trying to predict.

In many respects, commercial aviation already is a remarkable feat in uncertainty management. We routinely get on airplanes and reliably arrive at our destination, and fortunately our airplanes crash much less frequently than our computers. Airplanes move in the very complex system of the earth's atmosphere and together with air traffic control constitute one enormously complex system delivering remarkably reliable transportation. At small time and large scales the atmosphere is also turbulent and chaotic, and occasional crashes due to atmospheric disturbances remind us that this is not a triviality.

Because of this turbulence in the atmosphere and near the vehicle, there is chaotic dynamics surrounding the vehicle at every scale from the microscopic to the global. Furthermore, most objects having the size and mass of a 777 and traveling at high subsonic speeds would exhibit extremely unpredictable trajectories, although eventually hitting the ground at high velocities would be a certainty. Almost any other connection of the millions of parts in a 777 would also fail to behave predictably, although the 777 itself is remarkably robust to a wide variety of component failures. Despite tremendous advances in computation and its application to CFD and CAD, no simulations are ever performed that come close to capturing all this complexity at all these scales. Yet in spite of all of this, these millions of components manage to successfully "fly in formation" such as to deliver reliable and predictable performance. Fortunately, this success is not a mystical process (that the current antiquated air traffic control works at all *is* fairly astonishing), but it does involve tremendous amounts of domain-specific expertise and hand-crafted solutions. We must be realistic and cautious about the way in which VE technology should interact with this process.

VLSI Design

Design Today

Our third case history involves very large scale integration (VLSI)—a technology that enables millions of transistors to be fabricated onto a silicon chip less than a square inch in size. Such a chip can function as a complete microprocessor system performing hundreds of millions of arithmetic operations per second. For example, today's Intel's Pentium Pro microprocessor is faster than the most powerful supercomputers of the early 1980s, but costs 4 orders of magnitude less than those older machines. VLSI is an entirely different type of complex system than fluid dynamics, although both start with homogeneous materials: fluids and silicon/metal.

The number of transistors on a single chip, as well as the speed of microprocessors, has been roughly doubling every 18 months since the 1960s. This exponential-growth trend is known as "Moore's Law" (after Intel's co-founder

Gordon Moore) and is expected to continue at least until the year 2010. VLSI feature size (i.e., minimum width of a wire on a chip) has recently dropped to below one fifth of a micron (one millionth of a meter, or about one hundredth of the width of a human hair) and continues to steadily shrink into the deep submicron range. The next decade will usher in single-chip microprocessors containing hundreds of millions of transistors and operating at multi-giga-FLOPS (billions of floating-point operations per second) speeds.

Until the early 1970s, VLSI design was primarily done manually; such a labor-intensive process was possible because of the low gate-count (typically in the hundreds to low thousands) of those early circuits. However, when VLSI permitted fabrication of circuits with hundreds of thousands of gates, hand-crafted design was no longer practical (nor in many cases even possible). The field of computer-aided design (CAD) of VLSI circuits matured into a discipline with multiple annual technical conferences, and by the mid-1980s, dozens of commercial VLSI CAD software systems became available (typically at six-figures per copy).

A commercial VLSI CAD system is typically structured as follows. First, the circuit design is specified abstractly, with "blocks" representing high-level functional units (e.g., adders, multipliers, and memories); the general connectivity among these modules is also specified at that stage. Next, the design is further synthesized and specified in greater detail, with modules being fleshed out from an available library of predesigned parts (and some from scratch if necessary). If the design is too large to fit on a single chip, a "partitioning" step takes place, where the design is divided into several chip-sized modules (the typical optimization objective during partitioning is to minimize the wires that cross between modules, which in turn minimizes the number of wires that are forced to go off-chip and onto the circuit board, thereby slowing down the overall circuit). Next, a general "floorplan" is developed to accommodate the various circuit components on the chip area. Once a good placement is obtained for the circuit elements, the detailed interconnections are routed among the modules. Placement and routing typically attempt to minimize the overall chip area, which in turn minimizes the fabrication costs and thus maximizes the chip vendor's profit margin. Placement and routing also tend to be the most complex and time-consuming of the various VLSI design phases, since these problems are computationally intractable.

Extensive verification and testing occur at each level, with subsequent iteration and modification. These tests ensure that the circuit, once fabricated, will meet the design specification, both in terms of logical functionality and in terms of operating speed; the overall speed of the proposed circuit is determined by running a massive finite-element-based simulation over the circuit, which simulates the exact physical behavior of the circuit (such timing simulations typically take place at micron scales and at picosecond time resolutions—they often require days or weeks to run to completion). Despite such extensive testing, gross

bugs can still slip, especially when the error occurs at the higher levels of the design process (e.g., at the specification level); naturally, the longer a bug goes undetected, the greater the incurred cost to the company (e.g., the infamous Pentium FDIV bug cost Intel almost $1 billion to fix).

Overall, the field of VLSI design has evolved in the footsteps of the field of software engineering. Thus, it is not surprising that VLSI CAD embraces the basic precepts of hierarchical structured top-down design, functional orthogonality, component libraries (i.e., subroutines), design reuse, beta-testing, and so on. Unfortunately, as discussed below, by following the software engineering paradigm, the VLSI design process also inherits the problems inherent in that area.

The key feature of VLSI design that has traditionally greatly simplified uncertainty management is that the logic level could be effectively decoupled from the physical level. That is, through the use of electrical thresholding and Boolean logic, the chip design is constrained so that uncertainty at the microscopic level does not propagate to the macrolevel. Other techniques for uncertainty management in VLSI design include component redundancy and fault-tolerance algorithms, which automatically compensate for manufacturing defects and unforeseen system transients. One major advantage of digital representations generally is that error-correction codes enable perfect reconstruction and reproduction of the digital signal, so that small uncertainties have exactly zero impact on the final output. This is in stark contrast with turbulence where uncertainty at the microscopic level can easily have macroscopic effects. Thus, VLSI has a special place in complex systems, and this feature of digital systems has been emulated and exploited in other domains as well.

Future Trends in VLSI Design

The price paid for the complete isolation of component uncertainty in VLSI, and indeed the price paid for digital systems generally, is that system performance is sacrificed, sometimes enormously, in comparison with what would be possible if such an isolation were not maintained. This has been viewed as acceptable because it makes the design process manageable, and the evolution of fabrication technology allowed for both conservative design philosophies and dramatic progress in performance. Future trends in VLSI design may change this situation dramatically. The ever-shrinking feature size and growing gate counts of VLSI circuits are giving rise to a number of emerging trends and new fundamental problem areas.

First, thinner wires have a higher resistance, so interconnect delay now dominates the overall circuit speed (e.g., in current high-end designs, over three fourths of the clock period is due to signal propagation delays through wires). Moreover, wires take up most of the chip area, so VLSI physical design is now primarily a massive wiring-optimization problem (which is computationally intractable).

Typical problem instances are so large (and numerous), that even low-order polynomial-time algorithms are often much too slow in practical situations.

As feature sizes decrease into the ultra-submicron range, previously ignored phenomena become more significant, and indeed even begin to dominate the overall design. For example, parasitics such as capacitive-coupling between parallel wires (i.e., "cross-talk") has recently become a major problem since it tends to substantially increase signal propagation delay, as well as cause spurious signals/switching in the circuit. Another problem is "electro-migration," where the electrical current can randomly knock metal atoms out of the wires. This phenomenon tends to further "thin out" already-thin wires, which in turn exacerbates the electro-migration problem (i.e., creating a positive feedback loop), until open faults occur (i.e., wires become disconnected), which can disrupt the functional correctness of the overall circuit.

As feature sizes shrink even further, quantum mechanical phenomena will begin to have significant effects on circuit performance (e.g., quantum uncertainty, particle tunneling, and quantization of mass and charge). As clock frequencies advance into the gigahertz range, wires begin to behave like transmission lines, and parasitic phenomena such as signal attenuation and antenna effects become major performance-limiting concerns. The physical and electrical properties of most materials are still not sufficiently well understood at such high frequencies and small scales (e.g., current timing simulation techniques break down and are still not reliably applicable to regimes with feature sizes below one tenth of a micron).

Finally, at these small-scale regimes, even the VLSI manufacturing process itself becomes quite problematic. The confluence of these trends suggests that fundamental physics considerations will become a dominant component of VLSI design, the distinction between digital and analog circuits will become increasingly blurred (certainly, analog considerations will migrate up well into the higher levels of the VLSI design process), and uncertainty management will become a much more difficult problem.

COMPUTER SCIENCE, SOFTWARE ENGINEERING, COMPUTATIONAL COMPLEXITY, AND VE

In this section we touch upon a number of computer-related subjects that bear on VE.

Computer Science

A commonly shared dream is that in a VE environment, an engineer (or decision maker, analyst, commander, and so on) with a sophisticated virtual reality (VR) interface is connected to networks of other engineers in various disciplines sharing a common database. Design changes automatically propagate

so other engineers can respond by evaluating the consequences including manufacturability and maintenance. Immersive visualization methods take data from experiments or tests on physical prototypes and facilitate the comparison of data with theory and simulation. Engineering design of the manufacturing process and its product, together with training for operation and maintenance, can proceed simultaneously and synergistically. Much of the future promise of VE is based on tools from computation, workstations, supercomputers, networks, and software engineering.

The hope is that future VE software environments, including VR features, will relieve design engineers from the tedious, repetitive, and routine tasks that still dominate much of engineering design and let them focus on the critical decisions involving uncertainty management in the face of cost constraints. The fear is that it will also give highly unqualified people the illusion that they can click a few menu items, run some multidisciplinary optimization, and design a new airplane or automobile that will actually work in the real world. Sophisticated VE environments will inevitably increase the gap between the best engineers and the average, facilitating both the possibility of much better engineering and also the likelihood of spectacular failures.

The entertainment industry will push VR for an increasingly realistic look and feel and an emphasis on fooling human senses, and the many challenges in further developing VR are already well funded and appreciated. Although there will be many similarities in software design, the paradigm of "realistic = looks good" should not dominate VE as well. Otherwise, engineers and programmers will ultimately design systems fine-tuned for VR that do not work in reality. Without a correct and fundamental mathematical structure, VE could fail spectacularly and the potential for abusing it could be tremendous.

Many aspects of these issues are already well understood in the DOD M&S community. For example, advanced distributed simulation (ADS) and high-level architecture (HLA) are motivated in part by the recognition that high-fidelity engineering applications may involve timing or other issues too fine for human perception, and thus training-based systems such as distributed interactive simulation (DIS) are inadequate. Furthermore, it is well understood that all forms of M&S are sensitive to modeling and testing assumptions, poorly understood physical phenomena, or even to well-understood phenomena that require excessive computation. Nevertheless, it is still safe to say that the computer science view dominates the current VE landscape.

Software Engineering

Much of the current research in VE aims to make the design of complex systems more like the discipline of software engineering. To understand the implications of this, let us discuss briefly the history and basic trends in software engineering.

In the early days of computer science (1950s), computer hardware was expensive while computer software was free and quite primitive. By the mid-1960s, computer speed and memory capacity grew substantially, and along with computer performance grew the number of applications and user expectations. It was soon realized that building a large, complicated program is considerably more difficult than concatenating a series of smaller programs. "Programming in the large" often seemed to be an entirely new activity, exhibiting nonintuitive characteristics. To grapple with the growing complexity of software, more disciplined approaches to programming were explored; these included structured programming, strong type-checking, functional languages, program verification, software reuse, graphical user interfaces, and more recently, object-oriented design and programming, client-server applications, and network-based computing. However, none of these techniques proved to be a panacea, and each introduced new complexities and pitfalls.

Over the years, the problems associated with large software development efforts grew bigger and more pronounced. Some computer programs now contain over 15 million lines of code, and the programs for NASA's proposed space station will have more than 80 million lines of code. Despite modern programming tools such as interactive debuggers and visual programming environments, the average productivity of professional software engineers working on large systems is only a couple of dozen lines of code per day. Expensive and sometimes catastrophic system failures have been due in part to software bugs (e.g., the recent Ariane rocket explosion, the Denver airport baggage system, and lost NASA space probes). Many large software systems diverge so much from their planned project timelines and budgets that they are abandoned altogether, sometimes at a loss of billions of dollars. Two recent examples include the failed attempt by the FAA to revamp its aging air traffic control system and the aborted plan by the IRS to upgrade its software system for tax collection.

Why is it so difficult to write reliable code? A number of factors contribute to the difficulties. First, the complexity of hardware is quite bounded (i.e., it is only expected to be able to execute a relatively small and simple set of machine instructions), while the complexity of software is unbounded (i.e., software is expected to do *everything* else).

Second, it is very difficult to define or characterize the set of all possible types of inputs and conditions under which a system is expected to operate. This is particularly true in "embedded systems," computer hardware and software operating as part of a larger system, such as an airplane or automobile as opposed to a PC or network of workstations. If the user of a word processor does something the software does not expect, it can simply refuse to accept the input and the impact is minimal. If the control system in a launch vehicle (e.g., Ariane 5) receives some data that are unexpected, simply shutting down will (and did) result in huge losses.

Software is also inherently fragile and removing a single line of code could

render a large program completely useless or even dangerous. Contrast such instabilities with the robustness of living organisms, where the removal or death of a cell (or even many cells) typically has little or no impact on the overall functionality of the organism. Similarly, most complex engineering systems are deliberately designed to degrade gracefully under component failures.

Third, there is an inherent lack of logical symmetry in verifying and validating computer code: although it is easy to establish that a piece of code is buggy by simply exhibiting the particular error in question, proving the *absence* of bugs is usually impossible. That is, negative results are much more difficult to come by in logic and mathematics than positive results, and it is usually much easier to give a counter-example than to prove the nonexistence of something.

Fourth, humans are not very good at keeping track of thousands (let alone millions) of interacting parts, be it lines of code, transistors on a chip, or gears, levers, and pulleys. Short-term memory can typically store seven (plus or minus two) items, and perhaps a few additional items by utilizing some simple aggregation techniques. It is therefore not surprising that even with the aid of mechanical tools (i.e., mathematics and other formalisms), complicated systems with millions of interacting parts typically quickly diverge beyond our ability to understand them.

Fifth, the synchronization of effort among large groups of people working on a single, tightly coupled system becomes a logistical nightmare after some parallelism threshold is reached, and the communications load among the workers grows superlinearly.

Finally, in most human endeavors, the difference between the best and the average performance is a relatively small factor. For example, most people can comfortably run a 10-minute mile; on the other hand, a gold-medalist Olympic athlete can only outperform this mediocre record by less than a factor of three (similar performance ratios hold for other common skills, such as swimming, biking, jumping, lifting, typing, and reading). In contrast, the *best* programmers can be over a factor of 100 more effective and productive than the *average* programmer. Most code is by definition written by average programmers, and programming will always be a very labor-intensive activity; it is therefore crucial to select software team leaders and chief architects carefully.

In summary, the difficulties we encounter in software engineering are not unique to computer programs, but rather are fundamental to many areas of science and engineering that are concerned with building large, complicated systems (and indeed result from the limits of our current technology and our own inherently limited capabilities). We should therefore not expect "silver bullet" solutions anytime soon to the software engineering problem. Indeed, the broader VE environment is likely to exaggerate and magnify the problems of software engineering.

Having said all this, we need to also keep in mind that the field of software engineering (and computer science in general) has also made great advances.

Personal computers are more user-friendly, with elementary school children now routinely using computers and "surfing" the World Wide Web. Typical PCs are now faster (and have more memory) than the Cray I supercomputer of the late 1970s, yet cost under $2,000. Portable laptop computers, weighing around 5 pounds or less, have become ubiquitous. Interactive run-time environments, sophisticated debuggers, and visual programming languages have made basic programming easy to learn and to teach. Electronic mail has greatly facilitated communication and data exchange among people and researchers around the world, and the World Wide Web has made vast amounts of valuable information easily accessible to everyone. Thus, despite the various difficulties, software engineering has made great strides and contributions over the years as well.

Computational Complexity Theory

Since the beginning of the 20th century, a number of practical global optimization problems have been extensively studied. These problems include such classical formulations as Traveling Salesman, Boolean Satisfiability, Quadratic Programming, Hamiltonian Cycles, as well as a large variety of other partitioning, packing, placement, interconnection, routing, reachability, and approximation problems. Up until 1970, these problems had been attacked in isolation using ad hoc techniques, and in all cases researchers have failed to discover efficient (i.e., polynomial-time) algorithms for *any* of these problems. Nevertheless, no satisfying explanation existed as to why these problems all seem intractable, nor how these problems may be related to each other.

In the early 1970s, it was discovered that all of these problems are efficiently "reducible" to one another in a way that preserves solution quality (i.e., these algorithmic reductions map optimal solutions of any of these problems to optimal solutions of any other of these problems). This implies that if there existed an efficient algorithm for one of these problems, such an algorithm could be immediately (and mechanically) transformed into efficient algorithms for *all* of these problems. In other words, with respect to computational tractability, none of these problems is any more difficult than any of the other problems. Therefore, since none of these problems had been solved efficiently to that point, despite many decades of intense work by hundreds of good researchers, this unifying framework (technically referred to as "NP-completeness," and more generally as "computational complexity theory") provided the strongest evidence yet that all of these problems are computationally intractable. An NP-hard problem is classified as a problem that is at least as hard as the NP-complete class of problems.

These results provided a twofold contribution to the field of global optimization. First, once a computational problem is formally shown to be NP-hard, such a negative result saves much effort, since researchers need not bother to continue their search for an efficient algorithm to such a problem. Second, once we know that a problem is NP-hard, this gives us legitimate license to devise heuristic

approximate solutions, without having to worry that our work will be rendered obsolete any time soon by the discovery of an efficient exact algorithm. Today, these topics have become a mainstream subfield of computer science, with several yearly conferences being devoted to this subject.

Computational complexity theory is independent of what actual computers we use in practice and of our underlying computation model. This also means that speedups in VLSI technology or advances in parallel computing will not affect the class of problems that are "intractable." In short, computational complexity/intractability is fundamental, and cannot be overcome by "throwing silicon" at it.

Some problems are so intractable that there exist no algorithms whatsoever to solve them; such a problem is said to be "undecidable" (the topic of undecidability was pioneered by Alan Turing in the mid-1930s). Rather, it can be mathematically proved that for an undecidable problem, no algorithm exists whatsoever even in theory, no matter how complex or subtle a possible solution approach may be attempted (this is a much stronger negative result than the intractability/NP-completeness discussed above). Many undecidable problems are deceptively easy to state formally; for example, the problem of determining whether a given program runs forever (or halts eventually) over a given input is undecidable. In fact, any mathematical framework powerful enough to describe arbitrary programs (and this even includes simple arithmetic) is undecidable as well; such systems are said to be "computationally universal." This has very strong implications for VE, and for dynamical systems in general, since most of these systems are computationally universal (i.e., they can simulate arbitrary computations or computer programs, and are therefore undecidable). Thus, many interesting questions about dynamical systems (such as long-term behavior, quiescence, and termination) are undecidable, and there exist no algorithms for the resolution of these problems in general (although particular classes of simple instances may be solved in ad hoc ways).

Implications for Virtual Engineering

Both the history of software engineering and the theory of computational complexity have important implications for VE. While it is difficult to define or characterize the inputs to software, the uncertainty faced by complex systems is much greater than typically faced by most software. The inherent lack of logical symmetry in verifying computer code is even greater in more general complex systems. It is much easier to convincingly show that some design change will fail through simulation than it is to show that it will succeed. The former requires only one bad example, while the latter requires strong evidence of the complete absence of such examples. This is made even more severe in complex systems in uncertain environments.

As bad as we are at keeping track of million-line programs, and at working in

teams to write software, at least software is a highly homogeneous system and teams are usually composed of experts in at most two domains, software design and perhaps the application area that the software is targeted for. In VE we are faced with highly heterogeneous systems and teams. Also, the gap between the best and the average that shows up in software engineering is likely to be even greater in VE, and the consequences could be even more severe.

Paradoxically, many complex engineering systems work much more reliably than complex software systems. Thus, while we may expect VE to inherit many of the problems of software engineering, the constraints and discipline imposed by VE's connection with physical reality offer some differences with conventional software engineering that should not merely be overcome, but exploited. This underscores again the need for a theoretical foundation for VE that goes well beyond computer science.

Computational complexity theory also has a sobering message to deliver to a naively cheerful view of the future of VE. As we aim for cheaper, better, faster with complex systems for higher levels of performance, uncertainties in components and the environment will interact in new and unforeseen ways. Evaluating all the possibilities for failures due to these uncertainties is a computationally intractable problem, but one we cannot afford to ignore. A look back at famous failures of engineering systems will emphasize this point.

FAMOUS FAILURES OF COMPLEX ENGINEERING SYSTEMS

In this section we will briefly review case studies of famous failures of engineering systems: the Titanic, the Estonia Ferry sinking, the Tacoma Narrows Bridge collapse, subsynchronous resonance in power systems, telephone and power system outages, the Denver airport baggage handling system, and Ariane 5. While each of these failures was due partly or primarily to factors beyond engineering or technical considerations, we will concentrate on the technical issues. We have not included some of the most dramatic failures, such as Chernobyl, Challenger, or Bhopal, because these involve much more complicated interactions of engineering and human judgment, and they have received such extensive coverage.

We will argue later that there are unifying themes connecting these different disasters that are relevant to VE (dynamics, interconnection, and uncertainty management). We have suggested that complexity arises from the need to provide reliable predictability in the presence of uncertainty and that failures occur when uncertainties and interactions are not properly accounted for. These case studies will illustrate these issues and provide examples for a more extensive discussion in the next section.

In retrospect, for all of these failures, we can always identify a component that failed and do simple "back of the envelope" calculations with very simple models to explain the failure. It is essentially always possible to ignore, if we

choose, the system design issues that contributed to the failure. A deeper view also always reveals that there were system design flaws and that the apparent component failure was merely a symptom. Of course, the VE challenge is to create an environment where we are better at doing that *before* the failure occurs.[8]

Titanic

On April 14, 1912, the Titanic, the largest, most complex ship afloat, struck an iceberg and sank. It is generally agreed that the iceberg scraped along the starboard side of the ship, causing the plates to buckle and burst at the seams. Some investigators speculate that the ship was simply too large for the technology available; vibrations from its massive engines may have played some part in the buckling of the hull plates. The Titanic had a double-bottomed hull that was divided into 16 watertight compartments. Because four of these could be flooded without endangering the liner's buoyancy, it was considered unsinkable. Unfortunately, these compartments were not sealed off at the top, so water filled each compartment, tilting the ship, and then spilled over the top into the next one. Five compartments eventually flooded, slowly but surely sinking the ship. This is perhaps one of the all-time great failures to correctly model the interaction of uncertainty in the environment and the way it can couple with the dynamics of a system. A purely static view of the ship, one that ignored the dynamics of the water flow, would never have predicted the actual disaster.

Estonia Ferry

It would seem unlikely that a mistake of the type that occurred in the Titanic would be repeated. However, a weak door lock was one of the main reasons for the 1994 Estonia ferry disaster that caused the deaths of more than 800 people. The ferry's bow visor, a huge top-hinged door at the front of the ferry that swung up to allow vehicles to be driven into and out of the ferry's car deck was secured by several locks. The lower lock, known as the Atlantic lock, was too weak to withstand extremely heavy pounding by rough seas. Stormy seas in the Baltic Sea on September 28 broke the lock between 30 minutes and 1 hour before the 157-meter (515-foot) ferry sank shortly after midnight. The noise of the loose bow visor slamming at the hull was heard by several survivors. The slamming set off a chain of events, including the breaking of other locks, that ended in the tragedy. Only 137 of the more than 900 people on board survived. The commission that investigated the incident said the shipbuilder did not have proper blueprints for the lock when constructing the ferry in 1980. As a result, the commis-

[8]The following discussions abstract from many sources, written and verbal.

sion says the shipbuilder apparently made its own calculations and underestimated how strong the lock should be. This particular failure would seem the one most likely to be caught with an integrated CAD system.

Tacoma Narrows Bridge

The Tacoma Narrows Bridge was the first suspension bridge across the Narrows of Puget Sound, connecting the Olympic Peninsula with the mainland of Washington, and a landmark failure in engineering history. Four months after its opening, on the morning of November 7, 1940, in a wind of about 42 miles (68 km) per hour, the 2,800-foot (853-meter) main span went into a series of torsional oscillations, the amplitude of which steadily increased until the convolutions tore several suspenders loose, and the span broke up. The bridge was designed to have acceptable horizontal displacement under the static pressure of a much larger wind, but was not designed to handle the dynamic instability caused by an interaction of the winds and the high degree of flexibility of the light, narrow, two-lane bridge. Modeling this type of fluid-structure interaction, a particularly simple type of flutter, was within the technical capability of engineers at the time, but was evidently not considered. A modern analysis would likely view the fluid-structure flutter as a bifurcation problem, and analyze the nature of the bifurcation as the wind speed increased. Immediately after the accident, numerous investigators were able to create both simple mathematical and scale physical models that exhibited the same failure as the actual bridge, and very simple models were able to predict the wind speed that would cause the collapse.

Subsynchronous Resonance in Power Systems

Series capacitors are often used in AC transmission systems to provide impedance compensation, particularly for long lines with high inductance, at the 60-Hz synchronous transmission frequency. Series capacitors are economical ways to increase load-carrying capacity and enhance transient stability, but the capacitors can combine with the line inductance to create oscillators with natural frequencies below 60 Hz. These electrical oscillators can interact with mechanical torsional vibrational modes of the generator turbine shaft, and in some circumstances can cause instabilities that snap the shaft. This happened dramatically at the Mohave Generating Station in Southern Nevada in 1971 when the turbine shaft broke twice before the condition was properly diagnosed. This is a classic example of uncertainty management gone awry. The capacitors were introduced to improve the stability on the electrical side and reduce the potential vulnerability to electrical side disturbances, but they had the unanticipated effect of destabilizing the mechanical side. The phenomenon is now reasonably well understood and is taken very seriously in design of power systems.

Telephone and Power System Outages

In recent years, there has been an increasing rash of large-scale breakdowns of both the telephone and the power systems, typically triggered by small events that lead to a cascade of failures that eventually bring down large portions of the network. The high complexity and interconnectedness of these networks are designed to improve their performance and robustness, but can lead to extreme and unexpected sensitivity to small disturbances. In both cases, highly interconnected nationwide networks allow load balancing to be achieved more economically, and the resulting system is, in principle and usually in practice, much more robust to large disturbances or variations in demand. The high degree of connectivity also makes it possible for small failures to propagate and lead to massive outages. The solution to these sensitivities is to add additional complexity in the form of more sophisticated control strategies. Without careful design, this trend to increasing complexity will not improve robustness.

Denver Airport Baggage Handling System

The automated system was supposed to improve baggage handling by using a computer tracking system to direct baggage contained in unmanned carts that run on a track. Originally scheduled for completion in March 1994, the unfinished $234 million project helped postpone opening of the airport until February 1995. The delay reportedly cost the city roughly $1 million per day in operations costs and interest on bond issues, more than the direct cost of the project. Significant mechanical and software problems plagued the automated baggage handling system. In tests of the system, bags were misloaded, were misrouted, or fell out of telecarts, causing the system to jam. The baggage system continued to unload bags even though they were jammed on the conveyor belt, because the photo eye at this location could not detect the pile of bags on the belt and hence could not signal the system to stop. The baggage system also loaded bags into telecarts that were already full. Hence, some bags fell onto the tracks, again causing the telecarts to jam. This problem occurred because the system had lost track of which telecarts were loaded or unloaded during a previous jam. When the system came back online, it failed to show that the telecarts were loaded. The timing between the conveyor belts and the moving telecarts was not properly synchronized, causing bags to fall between the conveyor belt and the telecarts. The bags became wedged under the telecarts, which were bumping into each other near the load point.

Ariane 5

The Ariane 5 was not flight tested because there was so much confidence in the M&S. The first flight carried $500 million of satellites and was destroyed about 40 seconds after liftoff. The error that ultimately led to the destruction of

the Ariane 5 launcher was clearly identified in the report of the investigating committee: a program segment for converting a floating point number, representing a measurement, to a signed 16 bit integer was executed with an input data value outside the range representable by a signed 16 bit integer. This run time error (out of range, overflow), which arose in both the active and the backup computers at about the same time, was detected, and both computers shut themselves down. This resulted in the total loss of attitude control. The Ariane 5 turned uncontrollably, and aerodynamic forces broke the vehicle apart. This breakup was detected by an on-board monitor, which ignited the explosive charges to destroy the vehicle in the air. The code in question was reused from an earlier vehicle where the measurement would not have become large enough to cause this failure.

It is tempting to simply dismiss this as a software bug that would be eliminated by better software engineering. It is obvious that the programmer should have checked that the measurement was small enough that the conversion could take place, and if it could not, have the control system take some appropriate action rather than simply shut down. In this case the appropriate action would have been to do nothing, because this measurement, ironically, was not even needed after liftoff. This may seem to make it a trivial issue, but the same code did work fine on the Ariane 4, although a control engineer would presumably have preferred it be done differently.

While the "software bug" view has some truth, it is misleading, because the failure was due to dynamics of the Ariane 5 that were different from those of the Ariane 4. It is the interaction of the software with the uncertainty in the environment and the dynamics of the vehicle that caused the failure. This is not a software issue, but a design flaw at a much deeper level. It is likely the programmers responsible had no idea how to determine if the Ariane 5 had dynamics such that under suitable environmental conditions the measurement would be too large. Presumably, they could have consulted appropriate experts in control and aerodynamics and anticipated the problem, but it would not have been a computer science issue at all.

COMPLEX ENGINEERING SYSTEMS

General Observations

We have argued that, on the one hand, complexity is generally undesirable. It makes our models difficult to work with, and the case studies above suggest that it can lead to unexpected and disastrous failures. Yet we see an accelerating trend to build increasingly complex systems because uncertainty management demands that we introduce complexity in our models. Let us illustrate this now with two simple and familiar current examples: smart weapons and airbags.

Smart Weapons and Airbags

In smart weapons, sensors, actuators, and computers are added to counter uncertainties in atmospheric conditions, release conditions, and target movement. This yields reduced sensitivity to uncertainties in the environment, but at the price of increased sensitivity to a large number of new components. If a sensor or actuator component fails, the weapon may actually have much worse accuracy than a dumb weapon. If we are careful in our design, we can use this shift in vulnerability from uncertainty in the environment to uncertainty in our components to our great advantage by making sure that our critical components are sufficiently reliable. Interestingly, it could be argued that the most successful smart weapons so far have been the simplest, for example, Sidewinder and laser-guided bombs.

Automobile airbags also reduce vulnerability to uncertainties in the environment. With an airbag you are safer in a high-speed collision with, say, a drunk driver who has crossed into your lane. Since you have no direct control of the other driver's behavior, an airbag is one of the most cost-effective control strategies you can take. Unfortunately, there is again increased vulnerability to component failures. Even without component failures, airbags can make certain circumstances more dangerous. For example, a low-speed collision may cause the air bag to deploy even though without the airbag there would be no danger of injury. Thus one could be injured by the airbag itself under normal operation even when the system functions properly. This is particularly serious with small passengers, who may be in more danger with an airbag than without. Overall there is a substantial net reduction in fatalities, but increased danger of injury and death in certain circumstances for all people, and possibly a net increase in danger to smaller people.

The awareness of the danger of airbags to children and small adults has provoked a flurry of research to make more advanced and more complex airbags. Proposed schemes include making the airbag deployment more adaptable to individual differences in size and body position by using infrared and ultrasonic sensors, together with weight sensors and capacitance sensors, which detect water in human bodies. Unfortunately, it is possible to fool these sensors as bags of groceries with a hot pizza sitting on a wet towel could presumably be mistaken for a person. Lower-technology solutions include simply setting the threshold for airbag deployment higher so they go off less frequently in slower-speed collisions. All these solutions again highlight that the design is driven by uncertainty management, and complexity is introduced as a by-product.

What these two examples illustrate is a kind of conservation principle that is at work in complex systems. Indeed, as we will discuss later, control theory has several such conservation principles that are critical to understanding complex

systems. Informally, when we introduce new components to reduce the effects of uncertainty in the environment, we inevitably create increased vulnerability either to these new components, or to other uncertainties in the environment. Since we control the design, if we are careful we can use this tradeoff to our advantage and shift our vulnerability from things that are more uncertain to things that are less, but explicit models of uncertainty are critical in achieving this. Unfortunately, with increasing complexity, evaluating these tradeoffs can be conceptually and computationally overwhelming.

The earlier section on software engineering discussed how large software development projects require a highly structured approach throughout, since interconnection management dominates component design. While this is now and always will be a challenging domain, it is still relatively homogeneous domain with limited uncertainty. Complex systems engineering has all of the challenges of software engineering plus heterogeneity (hardware and software plus chemical, electrical, mechanical, fluid, communications, and so on) and greater uncertainty (in environment and in system components). Complex systems remain even more poorly understood than large software systems.

Complex systems are poorly understood in part simply because nonlinear, heterogeneous, interconnected, complex dynamical systems are intrinsically difficult to model and understand. But more importantly, the role of uncertainty is critical, but very poorly understood. Furthermore, scaling of problem size can make the interaction of these issues overwhelming. As we will see, control theory addresses uncertainty management explicitly, but from a very narrow perspective. A deeper understanding of complex systems is emerging, but in separate and fragmented technical disciplines.

Finally, there is the "referee effect." The referee effect comes from the observation that we notice referees only when they do a bad job. Similarly, we notice the details of our watches, televisions, phone systems, cars, planes, networks, and nuclear reactors only when they fail to provide reliable operation and shield us from the world's uncertainties. Basically, the product of a superior design process makes itself virtually invisible. Even when the design is flawed, it may appear to the user that the failure was due to some component, rather than an underlying design process. This is true in all the examples of failures above. Success or failure of components, including computer hardware and software, is relatively easily understood. The role of the system design process itself, deciding which components to use and how to interconnect them, remains a mystery outside of a narrow technical community. Thus complexity in engineering systems is very much in the eye of the beholder. A design engineer may deliberately introduce great complexity specifically for the purpose of providing the end user with an apparently simple and reliable system. The apparent complexity depends on the viewpoint, and traditionally the only global viewpoint is that of the control engineer.

LESSONS FROM CONTROLS

Increasingly complex systems rely on advanced control systems, from cheap, fast computer disk drives to fly-by-wire aircraft to automobiles, integrated chemical production complexes, semiconductor manufacturing systems, and manned and unmanned space systems. Yet, ironically, control engineering and theory remain poorly understood outside of a narrow technical community. Traditionally, control engineers have been responsible for system integration because the control engineer adds the last component to a complex system, and does systemwide uncertainty management. Generally speaking, however, control *theoreticians* generally do not support this process. The situation is changing dramatically, and the trend is to more integration of system design and control design, but we need to accelerate this trend, and control theorists must expand their vision and make greater contact with other disciplines.

Although control theory by itself offers only a piece of a potential foundation for a theory of VE, it provides a very important complement to dynamical systems and computer science because uncertainty management is the central issue in automatic control systems. The experience and successes and failures of control theory provide important technical foundation and additional insight into the potential role of theory in complex systems. Ironically, until the last 10 years, control theory and practical control engineering have had a very distant relationship. The old story was that since controls were the most mathematical part of engineering it should not be surprising that it simply took decades for theory to get from academia to practice. While this certainly has some truth, another view is that much of the theory was basically irrelevant, and the reason for this irrelevance was inadequate treatment of uncertainty.

Tremendous progress has occurred in just the last decade in developing a mathematical theory of analysis of *uncertain* systems in the subfield of robust control. The new tools of structured uncertainty, integral quadratic constraints, linear matrix inequalities, operator theoretic methods, and so on, are well beyond the scope of this appendix, but a few observations can be made. The rate of transition from theory to practice has increased dramatically, and ironically, control theorists are doing theory that is both more mathematical *and* more relevant. Another important factor is that they are using modern software tools to get their theory into CAD design packages that are commercially available. Thus theory is now routinely used in industry before it has had time to get through the review and journal publication process. The former can take months, while the latter still takes years.

One of the most important messages from control theory is that *there are fundamental conservation laws associated with uncertainty management in complex, interconnected systems*. The informal notion suggested by the smart weapon and airbag examples that vulnerability to uncertainty could not be absolutely reduced but could only be moved around has theoretical expression in the math-

ematics of control theory. There are conservation laws where the "conserved quantities" are related to net system-level robustness with respect to component and environmental uncertainty. Interestingly, some of these conservation laws (e.g., Bode's integral formula) are based on results that are up to 50 years old, although they are getting modern extensions. They do require upper division undergraduate mathematics to express, however, and are beyond the scope of this review. Like energy conservation, they limit the performance of interconnected systems, but with proper understanding can be manipulated to our advantage. Also, like energy conservation, attempts to violate them are constantly being attempted, often with catastrophic results.

While control theory must play a central role in a theory of VE, current control theory has many inadequacies that must be addressed in this broader context. The first and most obvious is that control theorists take a very limited view of system interconnection, assuming that there is a fixed "plant" with a well-defined performance objective and a controller with adequate sensors, actuators, and computation to achieve the performance. The control design then amounts to solving for the "control laws" that yield the desired performance. This view of control is no longer relevant to even today's design environment where the systemwide control engineer's view of performance is needed at the earliest design stages. As cost-effective uncertainty management correctly takes its place as the dominant design issue, control engineers are forced to play a broader role, and control theory must catch up just to address the current needs, let alone the expanded needs of future VE.

Another weakness of control theory is that it tends to treat uncertainty and nonlinearity completely separately. This has traditionally been a remarkably effective strategy. To illustrate this, consider the problem of reentry of the Shuttle orbiter. Viewed as a whole, the dynamics are extremely nonlinear, and there are substantial uncertainties. The strategy has traditionally been to use a simplified nonlinear model with no uncertainty to develop an idealized global trajectory for reentry, and then use a local linearized model to design a feedback controller to keep the vehicle close to the trajectory in the presence of uncertainty. The sources of uncertainty included atmospheric disturbances, unmodeled vehicle dynamics due primarily to unsteady aerodynamic and structural effects, parametric uncertainty in the mass distribution and aerodynamic coefficients, and nonlinearities. The nonlinearities include both those that were in the simplified global model, which have been eliminated through linearization, and also higher-order nonlinearities that were not represented even in the global model. Both are treated as sources of uncertainty in the linearized model. This strategy works well because the idealized trajectory creates a relative equilibrium about which a linearization is quite reasonable, and the effects of nonlinearities do not dominate the local behavior about the trajectories. It is easy to imagine many circumstances where this clean separation is not effective, because there is so much uncertainty that either the idealized trajectory is not meaningful or the local

behavior cannot be kept close enough to the idealized trajectory to allow the nonlinearities to be treated as uncertainties.

Control theory also has other weaknesses that must be overcome. While mathematical sophistication is a strength of control theorists, they must overcome the natural distance this tends to create with other engineering disciplines. This is one reason why control theory has been applied to dynamical systems and computational complexity with some early successes, but has achieved less success in other areas. The limited connection with modeling and physics is even more troubling, as control theorists tend to view modeling as a mystical and unpleasant activity to be performed by others, hopefully far away.

ANALYSIS OF UNCERTAIN DYNAMICAL SYSTEMS

While even a superficial exposition of the current state of the art in analysis of uncertain dynamical systems requires mathematics well beyond the scope of this paper, it is possible to suggest some of the ideas and difficulties with simple drawings. Recall the interference analysis. We can think of a three-dimensional solid component as being defined as a subset of real Euclidean 3-space. Thus, interference analysis is checking for any intersections of these subsets other than those that are specified. We can similarly think of components in a dynamical system as being defined as subsets of all the possible time trajectories that their state and boundary conditions can take. Thus, a circuit component can be thought of as specifying some set of currents and voltages, a mechanical component as specifying some set of velocities, positions, and forces, and so on. These sets are potentially very complicated as they are subsets of infinite dimensional spaces of time trajectories. Differential equations can be thought of as constraints that determine the set of behaviors.

An interconnection of components is equivalent to the intersection of the subsets that describe their behaviors. For example, two circuit elements connected at their terminals each constrains the signals between them, and an interconnection simply means that the constraints of both components are in effect. Engineering design may then be thought of as connecting components in such a way as to produce only a certain desired set of behaviors and no others. Undesirable behaviors are analogous to undesirable interferences in three-dimensional solids, in that they involve unwanted intersections of sets.

To make this point of view more concrete, recall the fluttering paper example, and assume we use a rigid body model of the paper in a case where the folds are fairly flat. The boundary conditions between the air and paper consist of the paper's position and orientation and their rates and the forces between the paper and the air. Both the paper and the air model put constraints on what these variables can be, and dropping the paper in air forces both sets of constraints to hold simultaneously. One solution consistent with the constraints is steady fall-

ing, but there are other fluttering motions that are also possible. The challenge in complex systems is discovering these extra solutions that may be undesirable.

If components are linear with no uncertainty, then their sets of behaviors are linear subspaces, and it is relatively easy to check globally for undesirable interconnections. This would be analogous to the three-dimensional solids all being just lines and planes. Uncertain or nonlinear components are more complicated to analyze. Very simple uncertain linear problems are NP hard, and simple nonlinear problems are undecidable. The strategy that has been exploited very successfully in robust control theory is a natural generalization of the bounding box idea to this setting of components of dynamical systems. Here the bounding boxes are in infinite dimensional spaces, and checking for their intersection requires sophisticated mathematical and computational machinery. So far, this is the only known method that successfully handles both parametric uncertainty and unmodeled dynamics and overcomes to some extent the intractability of these problems.

While the generalized bounding box methods (they are not called this in robust control theory, but are referred to with a variety of other, more technical terms) have been successful in control systems analysis and design (they are widely used throughout the world), their application outside of controls has been limited. What is particularly needed now is to put these methods more in the context of component interconnections, not just the plant-controller paradigm of standard control theory. Also, there remains a great need for methods to analyze uncertainty and nonlinearity together in some nontrivial way. Developing bifurcation analysis tools that allow for uncertainty would be a good initial step, and research in this direction is under way.

In robustness analysis of uncertain systems, it is usually much easier to find a failure if one exists than to guarantee that none exist when that is the case. This inherent asymmetry is present in three-dimensional interference analysis and software design and will be a major feature of VE. We must try to overcome this as much as possible, but recognize that a substantial asymmetry is unavoidable.

CASE STUDIES REVISITED

While we are far from having an integrated theory of VE, we can gather the various ideas we have discussed from dynamical systems, computer science, and control theory and briefly revisit the case studies. The success stories in the 777 solid modeling, in CFD, and in VLSI are encouraging, but extrapolation to the broader VE enterprise must be done with caution. Each success depends on very special features of the problem area, and there are substantial challenges within even these limited domains to extending the existing tools. None of these areas has faced up to uncertainty management in heterogeneous systems, though all are being increasingly faced with exactly that issue.

Among the failures considered, the Estonia Ferry disaster is the one most

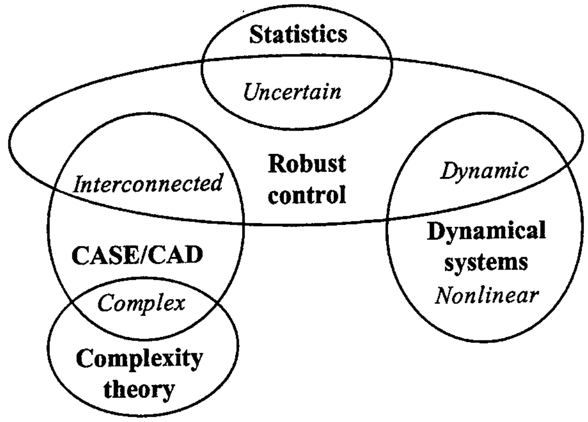

FIGURE B.11 Other foundations for VE theory.

likely to have benefited from the use of three-dimensional solid CAD tools such as were used for the 777. The Titanic, Tacoma Narrows Bridge, subsynchronous resonance, and Ariane 5 failures can all be traced to specific unmodeled dynamics whose analysis, had it been considered, was well within the capability available at the time. Thus it is easy after the fact to view these as simple problems with simple solutions, but the deeper question is whether a disciplined and systematic approach to VE would help avoid such mishaps. The answer is not obvious because each of these failures involved heterogeneous interactions and dynamics that are unlike the success stories.

The telephone and power system failures and the Denver airport baggage handling system fiasco are more clearly examples where uncertainty management in complex systems went awry. These highly interconnected and automated systems are intended to improve performance and robustness and at the same time reduce cost, and they generally do so with respect to the uncertainties and objectives that are considered primary in these systems. Unfortunately, the very complexity introduced to handle uncertainties in some aspects of the system's environment lead to vulnerabilities elsewhere.

It is tempting to imagine that a design environment that stressed uncertainty management and explicit representation of uncertainty across discipline boundaries would have encouraged design engineers to be alerted in advance to the potential for these failures, but we will have to wait until we have a better picture of exactly what such an environment would consist of. The challenge will be to avoid believing too much in either virtual worlds, or our past experiences with real ones, as both can mislead us about future realities.

Figure B.11 is intended to convey the way in which some existing communities are addressing the various aspects of VE models: uncertainty, interconnec-

tion, dynamics, nonlinearity, and complexity. It is intended to suggest that all the issues are being addressed, but in a fragmented way. We touched briefly and informally on all these topics except statistics. CASE here means computer-aided software engineering, and complexity theory is computational complexity in theoretical computer science. There are other areas that should contribute to a VE theory, such as nonequilibrium physics, all aspects of scientific computing and numerical methods, optimization, and discrete-event and hybrid systems.

We have argued that while sophisticated hardware and software infrastructures are needed to form the substrate on which robust VE tools can be implemented, the infrastructure aspects of M&S are already emphasized to a high degree, and the issues focused on in this appendix need comparable attention. In doing so we have perhaps paid inadequate attention to the need for new and novel software and user-interface paradigms that would address unique needs of VE. We regret we have had neither the time nor the expertise to explore this further. An aspect of computing that we will briefly discuss, since it is so ubiquitous, is so-called "soft computing."

SOFT AND HARD COMPUTING

Soft computing is usually taken to include fuzzy logic, neural-net computing, genetic algorithms, and so on, in contrast to the "hard computing techniques" of, say, numerical analysis, mathematical programming, structured and object-oriented programming, probability theory, differential equations, "hard" AI, and so on. According to its proponents, such as Lotfi Zadeh (see, e.g., Zadeh, 1994), "soft computing will revolutionize computing." While it is certainly beyond the scope of this appendix to give a thorough discussion of this area, we can provide at least one perspective. There are two standard arguments made for soft computing. The first is that many problems do not lend themselves to hard computing solutions because the systems under consideration are dominated by what we would traditionally call "soft" issues, like economic and societal systems, and anything involving human decision making, common-sense reasoning, and natural language. Hard computing and hard AI have failed to achieve long-standing goals of making human-computer interactions more human-friendly precisely because they have failed to appreciate soft computing approaches. Soft computing, especially fuzzy logic, allows programming with natural language.

Indeed, Zadeh has characterized fuzzy logic as "computing with words." The hope is that if you know a solution and you can simply and clearly articulate it in words, then you can program directly without translation to some programming language. There is substantial controversy regarding the degree to which fuzzy logic solves this problem, but the goal is certainly admirable, and there are cases in which fuzzy logic has been successful. On the other hand, many problems do not fit the fuzzy paradigm at all. In some cases, we can do a particular task well, but we cannot clearly articulate how we do it. Examples include chess

playing, vision, speech recognition, and almost all motor skills, such as those involved in sports or physical labor. Many of the tasks in which humans greatly outperform machines are also ones in which lower animals outperform humans. While biological systems do provide useful inspirations for machine automation, only humans typically articulate in words a detailed description of their own behavior. Perhaps more importantly, we often need the methods of mathematics, science, and engineering to help us find a solution. And in still other cases, using such methods permits us to find a better and more robust solution than possible with the simpler forms of fuzzy logic that have such intuitive appeal. By and large, we believe that the difficult problems of the VE enterprise are problems in which our naive intuition is likely to be dangerously wrong. In such cases, we should be cautious of seductive shortcuts.

The second argument for soft computing, and again fuzzy logic in particular, is that they more naturally exploit the tolerance for imprecision, uncertainty, partial truth, and approximation that characterize human reasoning.

In the context of VE, it is useful to distinguish two kinds of uncertainty:

1. The imprecision and ambiguities in our natural language, which parallels our sometimes limited ability to precisely specify what we want a system to do.
2. The uncertainty in our models of physical systems, as has been emphasized in this appendix.

While we have emphasized the latter, in the early stages in design of engineering systems the former can often dominate. If VE is successful in dealing effectively with type 2 uncertainty, then type 1 will be increasingly critical to overall system performance. It is here where fuzzy logic and soft computing hold the greatest promise. Advocates argue, though, that fuzzy logic is also ideally suited to handle uncertainty of type 2 as well. We disagree. Fuzzy logic is intended to capture properties of human language and simply does not address in any meaningful way many of the kinds of uncertainty we have discussed in this appendix and how uncertainty propagates with dynamics and interconnection. And, if one tried to use fuzzy logic to do so, it would quickly lose its comfortable "natural-language features." Fuzzy logic may be useful in representing human decision making in a simulation environment, but we have not considered that issue here. It may also be useful in a variety of engineering contexts that are ultimately much simpler than those in VE.

Similar remarks apply to genetic algorithms. Optimization, and particular global search techniques, will play a critical role in present and future VE systems. Indeed, our proto-VE examples of aircraft design with CFD, VLSI, and CAD of the type used in the Boeing 777 are domains where global optimization is either already playing a huge role (VLSI) or a growing role. Statistical methods, advanced optimization theory, and even theoretical computer science (decidability, NP-hardness) are creating a foundation for this subject, both in academic research

and in industrial application. From this point of view, genetic algorithms are a very minor piece of the picture. Their popularity is due primarily to the ease with which people can use them (people who include not only deeply capable scientists, but also more ordinary people with no or little expertise in statistics, optimization, or complexity theory).

Genetic algorithms are often mentioned as a moderately effective way to do global search, especially on highly unstructured problems. Based on our experience, which tends to be in hard areas of engineering rather than, say, softer problems of military combat modeling, we remain skeptical. Despite the strong market demands for commercial software to assist in global search in such problems as VLSI design and analysis of uncertain dynamical systems, genetic algorithms have had almost no impact relative to more mathematical approaches such as branch and bound, and problem-specific heuristics. This is not to say, however, that genetic algorithms have no role to play in VE. The conceptual simplicity of the approach means that it can be used by domain experts who may not be familiar with more sophisticated optimization ideas or may not want to invest the time to program a better algorithm. Genetic algorithms can be used to explore global optimization in a new domain, and if it is successful, then there is clear encouragement for further investigation. If not, little investment has been made.

"COMPLEX ADAPTIVE SYSTEMS" AND SOFT COMPLEXITY

A term that often arises in conjunction with soft computing is "complex adaptive systems," which can be considered to be a research area in its own right or a special case of what we have discussed here under the rubric of VE. It is not, however, a "new science," nor is it a substitute for the work we have described. Instead, what it has accomplished so far is to provide a set of metaphors for taking new looks at difficult problems involving complex systems. While the celebration of chaos, nonlinearity, and emergent phenomena has perhaps been overdone, and while popularizers have sometimes given them a nearly mystical flavor that seems bizarre to those of us working in the VE domain that includes control, dynamical systems, nonequilibrium physics, and complexity theory, the metaphors and popularized discussions have greatly broadened the audience and are helping to open minds regarding the value of experimenting with methods quite different from the traditional ones.

In this sense, work on complex adaptive systems is helpful to the VE enterprise. The concern, of course, is that the simplifications of popularization—which sometimes include exaggerated promises and claims—will discredit those associated with complexity research when those exaggerations are better recognized. This is a common problem in science. For example, there were backlashes against artificial intelligence and expert systems because the more exaggerated claims were finally recognized as such. The backlashes were sometimes quite unfortunate, because the research in these areas has had profound effects. In any

case, as we have indicated from the very beginning of this appendix, dynamical systems concepts will necessarily be at the very heart of any useful theory of VE.

It is important that VE researchers develop the kind of nonlinear intuition that the subject encourages and also build on existing methods for analysis of nonlinear systems. Both the concept of chaos—that apparent complexity and randomness can arise from deep simplicity—and the concept of emergence—that apparent order and simplicity can arise from deep complexity—are of great importance. On the other hand, they are empty without the more technical concepts such as phase space, bifurcation, strange attractors, Poincare maps, Lyapunov exponents, Hamiltonians, Euler-Lagrange equations, symplectic maps, integrability, self-organized criticality, ergodicity, and entropy. Unfortunately, there is no easy access to this deeper work.

To end this discussion, we might tentatively propose a notion of "soft complexity" analogous to, and including, "soft computing," in the same way that we might propose a notion of "hard complexity" that is analogous to and includes "hard computing." The flavor of the distinction would be as follows: Soft complexity equals emergence, fractals, artificial life, complex adaptive systems, edge of chaos, control of chaos, . . . plus soft computing, fuzzy logic, neural nets, and genetic algorithms. Hard complexity equals information theory, algorithmic complexity, computational complexity, dynamical systems, control theory, CASE/CAD, nonequilibrium physics, statistics, numerical analysis, and so on. This appendix has clearly advocated the relative importance of "hard" over "soft" complexity in VE. Some of the more extreme advocates for soft complexity claim it will revolutionize analysis and design of complex systems and obviate the need for the "structured and mathematical approach" advocated here. While we obviously disagree with this assessment, it is likely that soft complexity can help make concepts of hard complexity accessible, albeit in a limited way, to a nontechnical audience. It is also likely that the soft complexity concepts will be quite valuable in communication and, probably, for certain types of initial exploration of concepts. In any case, popular expositions of soft complexity will continue to emerge and will have effects on decisions about investment. Our hope is that papers such as the current appendix will help maintain perspectives.[9]

[9]For differing perspectives, see a selection of papers by users of fuzzy logic, including engineers, in *Proceedings of the IEEE,* March 1995. See also the collections of Zadeh's papers (Yager et al., 1987). And, in this volume, see Appendix G for examples of fuzzy logic research.

C

Simulation-based Acquisition

Richard Ivanetich, Institute for Defense Analyses

INTRODUCTION

The objective of simulation-based acquisition (SBA) is to enable the acquisition process to proceed in a highly integrated and collaborative manner. Any military system represents the interests of numerous parties—e.g., its operators, acquisition authority, designers, producers, and maintainers. The integration envisioned by SBA allows all these parties to interact closely during the development of the system so that the resultant system reflects as well as possible their combined interests, the necessary tradeoffs between their individual interests having been reconciled in an optimal manner from the overall perspective of the system. In this way, the highest-quality system at the least cost should be obtainable. The opposite extreme is a "stovepipe" process, where the interests of the "downstream" communities (e.g., maintainers) are not represented adequately in the "up-front" design of the system, or such interests are later accommodated by expensive modifications to the system.

Integration and collaboration do exist in the acquisition process today, but the intent of SBA is to extend this capability greatly. From a technical perspective, achievement of this capability centers on the concept of a collaborative environment of design, analysis, and simulation tools in which a computer-based representation of the system under consideration (e.g., ship, aircraft, or major component of such)—a so-called virtual prototype—may be built and examined. This virtual prototype will be available to all concerned with the system, and they may examine it using their design, analysis, and simulation tools. The fact that all parties deal with and have ready access to a common representation of the

system is the key element enabling these parties to interact in a highly integrated and collaborative manner.

These concepts are explored further in the following sections, beginning first with a more detailed discussion of the use of SBA.

SBA IN THE LIFE-CYCLE PROCESS

The life-cycle phases of a system may be specified progressively as follows:

- Requirements definition,
- Concept exploration—different concepts to meet the requirements are explored at a high level, and one option is chosen,
- Engineering design—the chosen high-level design is converted into a detailed design suitable for production,
- Manufacture,
- Test and evaluation—both developmental and operational, and
- Operation and maintenance—includes training necessary for operation.

System upgrade proceeds through these same phases, too, although the first five might not be as extensive compared to the case for a new system.

SBA relates to these phases in two ways. First, the concerns across the life cycle can be explored early in the life cycle. For example,

- The high-level virtual prototypes developed during concept exploration can be examined by the operators and maintainers to see how well the proposed concepts will meet their needs, with suggestions for improvement being fed back into the concept exploration. This assessment can be accomplished by visual examination of physical configurations depicted by virtual prototypes and by exercising the virtual prototypes in combat simulations.
- The engineering designers can comment on aspects of the high-level design that would be particularly expensive to realize, and discussion initiated with the requirements developers to see if less costly tradeoffs can be made.
- The detailed engineering designs can be examined by the manufacturers for production feasibility and suggested changes in design to simplify production processes.

Second, because of the integration afforded by SBA, design and other products developed during a life-cycle phase can be passed on to the next phase, thereby ensuring greater continuity and allowing cost reduction through reuse. For example:

- Engineering design can begin as a natural extension of the high-level design.
- The detailed engineering design can automatically be used to calibrate manufacturing processes.

• Simulations developed for conceptual exploration can be used for training purposes in the operations phase.

In short, the virtual prototypes and their associated design representations can serve as a record that is passed across the system life cycle.[1]

SBA IN THE ENGINEERING PROCESS

Just as SBA promotes integration and collaboration across all life-cycle phases, it also promotes such within each phase. One phase that should be particularly noted in this regard is engineering design because it can involve very large design teams representing many engineering disciplines. Activity both within and across the disciplines must be coordinated. SBA seeks to facilitate the flow of information across the design team and increase the ability of its members to readily access this information. Consequent benefits would be as follows:

• The design time will be shortened because of the increased support in design tools and the much more ready access to design information.
• The decreased design time will allow more detailed design options to be considered, thereby leading to a better and possibly less costly design.
• A more nearly optimal design can be achieved by optimizing the design simultaneously from the perspective of all disciplines involved (e.g., aerodynamics, structures, and materials in aircraft design), rather than by suboptimizations conducted one discipline at a time.

Furthermore, if significant changes in design or tradeoffs not anticipated during concept exploration are introduced, the associated virtual prototype can be fed back to other members of the overall community (e.g., operators, maintainers) so they can assess and comment as necessary upon the effect of the changes.

STATUS OF SBA TECHNICAL CAPABILITIES

Three key components are necessary to provide the overall SBA collaborative environment:

• Product representations—these are the computer-based representations of a system or the components of a system. They should refer to both the design of

[1]The virtual prototypes will evolve over the life cycle. The high-level ones developed in concept exploration will become more detailed during engineering design; as hardware components become available, they could for certain purposes replace simulated ones in the virtual prototypes; and the data used in the virtual prototypes will be refined based on information gathered in test and evaluation.

the system or component and its behavior. The behavior is determined by a model of the system or component. The model could be something relatively straightforward like specifying the weight of a beam used as a structural component in a ship or something more complex like the performance of the overall ship. A product representation capable of describing behavior is said to be an executable product representation.

- Analysis tools—these are any of the large number of tools that would be used, for example, in creating or assessing the designs and associated product representations. Computer-aided design (CAD) tools are a primary example.
- Interface infrastructure—this is the capability that allows the product representations and tools to interact with one another.

The following sections briefly describe the status of capabilities in each of these areas, in terms of both existing capabilities and missing capabilities.[2]

Product Representations

Existing Capabilities

Sophisticated static (nonexecutable) representations of the geometric aspects of a design exist. This has been demonstrated, for example, in the development of the Boeing 777 aircraft, where all the data describing the physical configuration of the aircraft were generated, manipulated, and retained in digital form accessible to all the designers. Furthermore, sophisticated executable representations of structural components (e.g., beams, bulkhead plates) have been demonstrated and used, for instance, in the development of tankers by Newport News Shipbuilding. In this case, for example, the dimensions of the structural component can be changed in the digital representation by the designer, and the properties of the component (e.g., weight) are automatically recomputed.

Missing Capabilities

Three related missing capabilities have been identified:

- *Multi-resolution modeling formalism.* Typically, one wants to predict the performance of an overall system from the properties of its components. This has been done manually, but there exists no formalism to facilitate the ready aggregation or disaggregation of product behavior.

[2]Not included here is a discussion of the combat simulations in which the virtual prototypes would be exercised. Such simulations are a key component of SBA, but are a major topic of discussion elsewhere in this report.

• *Cross-domain consistency.* It is unrealistic to think that there will be just one "object" describing the behavior of a system or component. Members of different disciplines will have their own representations expressing those properties of interest to them (e.g., some will be interested in thermal properties; others will not). A formalism is necessary to help ensure that these different representations are consistent (e.g., mean the same thing by a commonly named variable).

• *Propagation of uncertainty.* No matter how detailed, there is always some element of uncertainty in the description of a system or component. Individual disciplines (e.g., aerodynamics, structural mechanics) have characterized these uncertainties fairly well. What has not been treated, however, is how uncertainties in the model of one discipline propagate when the model is used in conjunction with the model of another discipline. This is necessary to understand because the overall behavior of a system or component is predicted based on these multiple models.

Analysis Tools

Numerous sophisticated tools (e.g., CAD) are available in modern engineering environments and will not be detailed further here. However, full realization of SBA does require further capabilities, two particularly critical ones being design optimization and cost estimation tools. In particular, multidisciplinary design optimization methods and tools—which allow attempts at design optimization to proceed across all disciplines concurrently—are necessary. Such methods have been illustrated in simple examples (e.g., Advanced Surface Combatant demonstration in the DARPA Simulation-based Design (SBD) program), but development of sophisticated, comprehensive techniques has not yet been achieved. Cost estimates of designs are possible, but there do not exist cost models that easily predict the results of changes in design parameters, especially in cases where new technologies are involved for the system or component under consideration. Such capabilities are necessary to readily effect the cost-performance tradeoff analyses envisioned as an aspect of SBA.

Interface Infrastructure

Existing Capabilities

The general concept of integrating product representations and tools has been demonstrated by Lockheed-Martin and other contractors as part of the SBD program. A key feature here was the ability to put software "wrappers" around existing simulations and tools so that they could interact with one another. In this way, it is possible to use existing components to conduct SBA, and not necessarily require the construction of new components.

Given the wrapped components, the basic interfaces among them were de-

fined by the high-level architecture (HLA) for simulation, although it was also necessary to define domain specific interfaces for the applications considered. The HLA object model templates also provided the basis for defining the nature of the information to be exchanged among the simulations and design tools. Use of HLA was demonstrated in the engineering proto-federation experiment carried out as part of the HLA program, as well as in the DARPA SBD program.

Missing Capabilities

As noted, it was necessary to define domain-specific interfaces to carry out the SBA demonstrations. In general, it will be necessary to define standards for such interfaces so that SBA product representations and tools can be developed independently and shared. No such standards now exist. For instance, important examples relate to geometric modeling representations. No standards now exist to allow the ready coupling of these data to dynamics simulations used for design purposes (e.g., computation fluid dynamic calculations) or to numerically controlled manufacturing devices.

SUMMARY OF SBA TECHNICAL CAPABILITIES

A significant SBA capability now exists, particularly as relates to the use of shared digital representations for depicting physical configurations in place of "paper" representations. For example, [3,4,5]

- Errors in design can be detected and corrected much earlier, as evidenced by Boeing's 95 percent reduction in engineering change notices in going from the 757 to the 777 aircraft. Similarly, rework on produced aircraft was reduced from 30 percent on the 747 to 3 percent on the 777.
- The design cycle time can be significantly shortened, as evidenced by the 20 percent reduction in cycle time achieved by Newport News Shipbuilding in developing the Double Eagle tanker.
- Significant cost reductions should also be achievable, as evidenced by the projected 25 to 30 percent cost schedule reduction due to elimination of the physical mockup in development of the NSSN by Electric Boat.

While the above examples refer primarily to the use of geometric modeling and executable product representations for structural components, a demonstra-

[3]Visit to Boeing Corporation, March 14, 1997.

[4]"Simulation-based Design" briefing at Lockheed-Martin, February 28, 1997.

[5]See Patenaude (1996). This report (conducted for the Deputy Director, Test, Systems Engineering and Evaluation, Office of the Secretary of Defense) contains several examples of the use of modeling and simulation in the acquisition process.

tion illustrating the more general concept of SBA has been conducted in the SBD program. In that example, design tools, engineering simulations, and combat simulations were integrated, operating on a common product representation.

Still, significant additional technical capability as noted in the subsections above, is required to achieve the full SBA capability. The multi-resolution modeling formalism will allow one to move more readily from component to system representations, as is necessary to exercise engineering-level designs in combat simulations. Multidisciplinary optimization, cross-domain consistency, and the propagation of uncertainty are all necessary to achieve the full degree of collaboration envisioned by SBA for the engineering process. And domain-specific interface standards are required to allow the ready integration of SBA components developed independently by different parties. Some of these needed additional capabilities such as the interface standards might be achievable by disciplined coordination efforts in the SBA community, but most of the additional capabilities are still at the status of difficult research problems today.

CULTURAL FACTORS

SBA is not solely a technical matter. In fact, since it presents new methods for the acquisition process, there are also factors of cultural and managerial acceptance. In some regards, these factors could provide challenges as significant as the technical ones. There appears to be a growing acceptance of SBA in commercial industry—at least from the perspective of geometric modeling and executable representations of structural components. For example, the Boeing and Newport News experiences were noted above, and Lockheed-Martin is also applying SBA in a satellite development program. Within DOD, significant interest has been expressed in OASN(RDA)/ARO and OSD/DOT&E, as well as in the DARPA SBD program. Individual Navy programs (e.g., SC-21, CVX, NSSN, LPD-17) have also expressed some interest. However, no institutional commitment to SBA has been made by the Navy.

One factor relating to the acceptance of SBA is that it requires greater up-front cost, although at the promise of significantly reduced life-cycle cost and possible reuse in other programs. These up-front costs relate to the development of greater design artifacts and simulations. Program managers can be reluctant to incur these costs because their benefits will not be realized during the program manager's tenure. A higher-level institutional commitment could thus be required to promote SBA.

CONCLUSIONS

The capabilities envisioned for SBA offer significant potential for providing a more efficient and effective acquisition process. In part, these capabilities have been realized and their benefit shown in examples such as the Boeing and New-

port News ones noted above. Furthermore, the overall concepts have been illustrated in the SBD demonstrations. Still, as noted above, significant technical and cultural challenges remain before the vision of SBA is obtained. What follows indicates steps by which the Department of the Navy, and DOD more generally, can work toward this vision. It should be emphasized that the path to the full vision is long and complex enough that it is not adequate just to postulate this vision. Rather, any planning done in the Department of the Navy and DOD should lay out a logical set of steps to this vision providing increasing capability, and also assess the cost-effectiveness of each of these steps. The following actions would provide some of these steps:

- *Pilot projects.* As noted, program managers could be reluctant to institute SBA capabilities because of up-front costs to their program, even though there could be significant downstream benefits. Thus, a separate pilot project or projects could be set up to develop SBA capabilities that would feed into a major naval program (e.g., ship or aircraft). Examples of the sort of capabilities developed would be (1) executable product representations (in particular, ones that go beyond the current structural representations), (2) the coupling of representations to combat simulations to assess the utility of the designs being created, (3) the coupling of geometric representation data to dynamic engineering simulations, and (4) means to share information (in both directions) between design engineers and manufacturing producers so as to both enhance the producibility of designs and also let the designs take advantage of new manufacturing concepts. In addition to helping an individual naval program, a pilot project would also promote demonstration, assessment, and transition of SBA capability in the Navy more generally.
- *Standards development.* While the HLA defines some general interfaces for integrating simulations, more extensive domain-specific interface standards are necessary to allow independently developed design tools, product representations, and simulations to interact with one another. The degree of standardization is not clear a priori. Not all aspects of the interfaces should be standardized, less the standards become too constraining, but some core set should be standardized. Experimentation would be conducted to determine this core set. Such experimentation would be carried out by getting together a set of participants representing members of the life-cycle and engineering communities and letting them work out the standards (including associated semantics) in the context of a demonstration project. This approach would be analogous to the proto-federation experiments carried out under the direction of DMSO in HLA development.
- *Research.* Even with the items noted in the previous two paragraphs, some significant capabilities requiring basic research will also be necessary to achieve the full SBA vision. Relevant topics include multi-resolution modeling formalism, cross-domain consistency, propagation of uncertainty, multidisciplinary optimization, and advanced cost estimating tools. Research programs are

needed to address these problems. They should be conducted in coordination with parties in the other Services and organizations like DARPA and DMSO that would also be interested in these matters. While one cannot predict just when basic research will have fruitful results, one should attempt to guide the researchers by having them apply their results to concrete problems as soon as feasible.

The costs associated with such steps can only be roughly estimated here. A pilot program might cost around $20 million to $50 million per year and run for 2 to 3 years. The experimentation to determine standards might cost approximately $20 million per year and run for 2 years. The research program might cost on the order of $10 million to $20 million per year and run for several years.

D

Exploratory Analysis

Paul K. Davis, RAND and the RAND Graduate School

In Chapter 4 of this report, it is argued that most models used to describe phenomena relevant to military operations, training, or acquisition will contain substantial uncertainty. This uncertainty can arise, for example, from a lack of knowledge about the operational circumstances of future battles, the combat processes being described, simplifying assumptions that lead to stochastic components in the model, or human behavioral elements. While such uncertainty is generally intrinsic to such models, all too frequently attempts are made to remove uncertainty from the model, that is, to suppress the issue. For example, stochastic effects are replaced by a notion of their average value. Parameter estimates of highly uncertain variables (e.g., a future war's warning time) are treated as correct. Moreover, if uncertainty is recognized at all, it usually is through conducting sensitivity analyses on a few variables while pretending that other highly uncertain variables are known. While such an approach can often be seriously misleading, it is difficult indeed to treat uncertainty comprehensively. Techniques such as exploratory analysis are just now becoming increasingly available; the difficulties cannot be underestimated, and considerable research on this problem will be needed for years. This should include development of new analytical tools.

Exploratory analysis attempts to seriously confront uncertainty in a given model rather than ignoring or removing it. When uncertainty is involved, however, the parameter space in "soft problems" such as those that arise when considering operational-level planning becomes very large. One possible approach is to run the model over a huge domain of parameter values (input assumptions)—not merely in the manner of common sensitivity analysis, but in ways that examine

much of the outcome space.[1] Of course, there are temporal limitations to this approach—even more so if some of the parameters are stochastic, requiring repeated runs to establish a distribution of results. Even without stochastic effects, problem dimensionality can explode as one acknowledges additional uncertainty. As a result, it becomes necessary to adopt a highly structured approach. For example, the field of statistical design of experiments with tools such as fractional factorial or Latin hypercube designs can substantially reduce the number of trials needed to identify important variables, significant combinations of variables, and the optimal combination of variables.[2]

These design techniques are widely used in industrial applications (albeit applications with fewer uncertain variables than often occur in military problems) where the purpose is to determine the best combination of variables to optimize an industrial process. Generally, a fractional factorial design identifies a relatively small number of experiments to be run of a highly structured sort. Once the results from these runs have been obtained, some variables are identified as being important, and a new set of runs is determined. This process continues as long as time and resources permit. At the end, one obtains reliable information on the most significant variables or combinations of variables and their influence on the outcome. As computing power increases, the size and complexity of problems that can be explored in the fashion will also increase. Thus, statistical design of experiments holds the promise of being an approach to cope with uncertainties in complex models. Much progress has been made (Bos, et al., 1978; Davis, 1994), but much more work needs to be done to tailor these methods to problems of military relevance.[3]

This approach represents a sharp departure from the long-standing legacy of using allegedly representative "point scenarios" and altogether ignoring major uncertainties (e.g., regarding the fighting capability, for constant equipment, of different nations' forces, or the "true" equation describing the movement rate of a division as a function of various combat variables).

Unfortunately, current M&S has not been designed with uncertainty analysis

[1]RAND has done considerable work on this approach over the last decade, beginning with development of the RAND Strategy Assessment System (RSAS), which evolved into the JICM operational-level model, sponsored by OSD's Director of Net Assessment (see Davis and Winnefeld, 1983, pp. 62-65 for early visions). The original technology, however, was not yet powerful enough for what is becoming feasible now. For a broad and thorough description of exploratory modeling and analysis from a computer science perspective, see Bankes (1993, 1996). For applications to defense planning and adaptive planning involving global warming, see Davis et al. (1996) and Lempert et al. (1996) (a reprint from the journal article in *Climatic Change*, 33(2), 1996).

[2]For practical discussion of such matters and citations to the literature on experimental design, see Committee on National Statistics (1995).

[3]Alternative approaches or formulations are possible as well. For example, control theorists have focused on consequences of unmodeled dynamics. Dynamical systems focus on chaos as the explanation for apparent random behaviors. These are discussed in Appendix B.

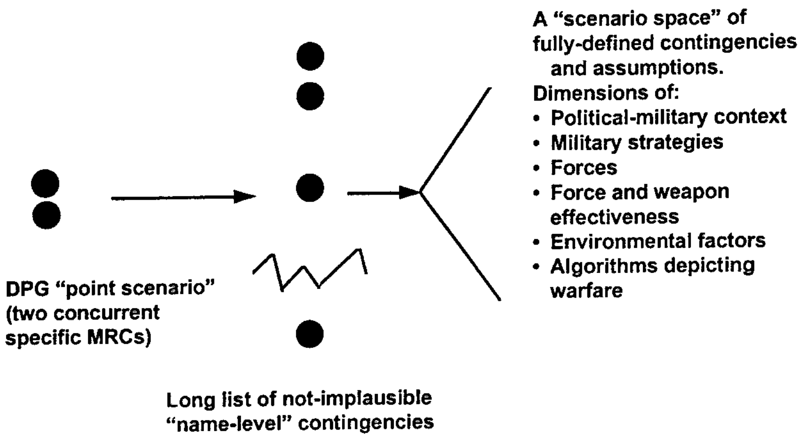

FIGURE D.1 Moving from point scenarios to a scenario-space exploration.

of this sort in mind. Yes, some M&S accommodates varying some types of data readily, but almost no M&S has yet been designed to facilitate meaningful exploratory analysis of the sort we have in mind here. Much less has it been designed with tools permitting workers to search for critical domains or draw synoptic conclusions. Much could be done, however, with future M&S, assuming appropriate designs and infrastructure. One key to such design is multi-resolution modeling, as discussed in Appendix E. Such an emphasis on uncertainty analysis would revolutionize the use of "soft" models such as those describing force-on-force battles and operational and theater-level conflict. It would also be essential for the engineering of complex physical systems that must operate in diverse circumstances of environment, tempo, and commander style.

Figure D.1 illustrates the concept in the context of rethinking higher-level defense planning. It depicts moving from point scenarios such as those used in the Defense Planning Guidance to an exploratory analysis framework. It shows expanding the set of "name-level" scenarios, and then recognizing that each such scenario (e.g., Iraq versus Kuwait) actually consists of an infinite number of variations. These can be explored by conceiving the "scenario space" formed by the axes shown: political-military context (e.g., who is allied with whom, what are the objectives, and what are the time lines); military strategies; forces; force and weapon effectiveness (remembering that planning factors are often wrong); environmental factors (e.g., weather); and, finally, the algorithms and algorithm parameters depicting warfare (despite pretenses to the contrary, these are highly uncertain as well).

Having conceived the scenario space, one can—in the context of a particular study—design an exploratory analysis covering the uncertainties of interest. One can then use modern computers and graphics to "fly through the outcome space"

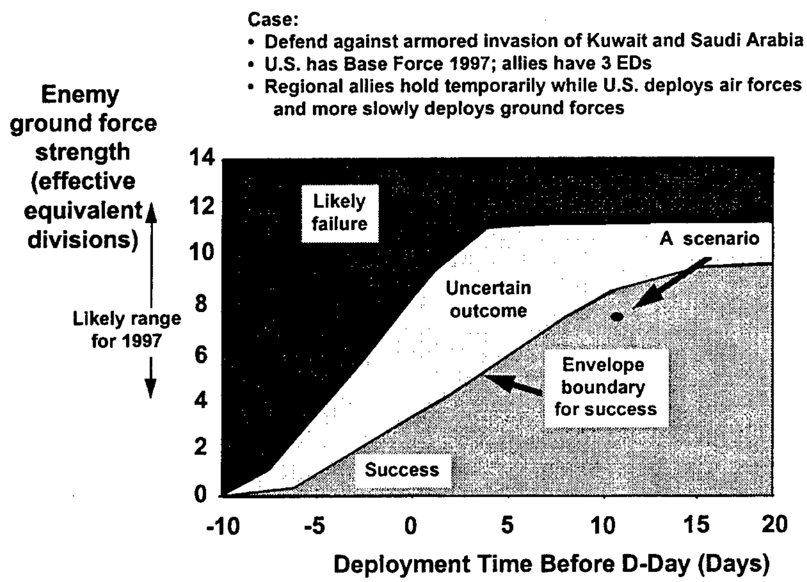

FIGURE D.2 Capabilities through a slice of scenario space.

to see when (under what assumptions) war outcomes would be favorable, unfavorable, and so on (Figure D.2). The purpose, of course, is insight. With enough insight, one could do a great deal to hedge against ever being in one of the "bad" regions. Some of the hedges would be obvious (e.g., prepositioning to increase deployment rates), but others might be less so (e.g., having a variety of systems to avoid common-mode failures of critical precision-strike weapons)—until after the exploration makes them obvious.[4]

[4]See Davis, Gompert, and Kugler (1996) for a relatively short account of this work and some of the unclassified insights from initial exploratory analysis of future regional contingencies, the upshot of which was to focus attention on Achilles' heel problems and the potential for "asymmetric strategies" by the adversary, rather than different ways to add marginally to the already substantial U.S. capability for "canonical" major regional contingencies with, for example, good use of warning and effective allies. For more discussion of how this relates to adaptive planning for military operations, see "Planning for Adaptiveness" in Davis (1994), which summarizes work over the preceding half-dozen years. A number of the ideas and methods referred to in this work were applied in the 1997 Quadrennial Defense Review. For an independent discussion of similar ideas, see the work of Bonder and Cherry in Vector Research, Inc. (1992).

E

Multi-resolution Modeling and Integrated Families of Models

Paul K. Davis, RAND and the RAND Graduate School
Bernard Zeigler, University of Arizona

INTRODUCTION

This appendix discusses multi-resolution modeling (MRM) and the related subject of integrated families of models.[1] These have to do with changing resolution within a single model or connecting two or more models and—the key issue—doing so in a substantively valid way.

Reasons for Interest in MRM

The reasons for wanting multi-resolution modeling are many, but they relate ultimately to the fact that we interact with the world at many different levels of resolution. We depend on low-resolution for (1) making initial cuts at problems, (2) "comprehending the whole" without being lost in the trees, (3) reasoning about issues quickly, (4) analyzing choices in the presence of uncertainty, (5) using low-resolution information, and (6) helping to calibrate higher-resolution models.

We also need high-resolution models for many purposes, notably (1) to understand underlying phenomena, (2) to represent and reason about detailed knowledge, (3) to simulate "reality" and create virtual laboratories for studying

[1]This appendix is largely based on work reported in Davis and Huber (1992), Davis (1993), and a review article discussing a related conference (Davis and Hillestad, 1993a,b). Some other material is adapted from presentations at a minisymposium, "Linking Simulations for Analysis," held by the Military Operations Research Society (MORS) in Albuquerque, N.Mex., February 25-26, 1997. The appendix also reflects discussions with Ben Wise, Paul F. Reynolds and students at the University of Virginia, Richard Hillestad of RAND, and Judith Dahmann of the Defense Modeling and Simulation Office (DMSO).

phenomena that cannot be studied in any other way (e.g., a range of possible battles and wars), (4) to use high-resolution information, which is sometimes quite tangible (e.g., weapon performance), and (5) to help calibrate lower-resolution models.

This need for models at different levels of resolution will not change merely because computers become more capable. Thus, we also need to understand the relationships among phenomena at the different levels, which in practice means understanding how models at those levels should relate to each other.

Reasons for Interest in Connecting Models of Different Resolution

It is also necessary to *connect* models of different resolution. Connections may be in software, so that one model takes data electronically from another, or "offline" (by what is humorously known as "sneakerware"), where humans take data from one model and then feed it to another, often massaging it during the transfer.

If the only purposes were analytical, then it might be sufficient and desirable to work with model families—when good ones existed. From time to time, one would cross-calibrate the models to ensure consistency with all known information. Most of the time, however, one would use a specific model tailored to the problem.

With the advent of distributed simulation, however, much is changing. The need now exists to connect a variety of models, often with different resolutions, and to do so at run time. Further, as computing power has increased, some workers have become interested in doing analysis with models that normally operate at one level of resolution, but occasionally call higher-resolution subroutines.

There are many reasons for operating at multiple levels in an advanced distributed simulation (ADS) environment. One objective is to avoid high resolution except when needed with the purposes of (1) conserving network and CPU resources; (2) simplifying and accelerating scenario setup; (3) reducing the number of simulation operators; (4) speeding simulation execution; and (5) simplifying setup and execution of low-priority "context" segments of a simulation while allowing detailed and authoritative representation of high-priority segments. Another purpose is connecting legacy simulations written at different levels of resolution.

Scope of the Challenge

Table E.1 reminds us of the basic levels at which military issues must be studied.[2] Work at these levels requires different models, but a planner at one level (e.g., a joint task force (JTF) commander) cares whether his planning frame-

[2] Adapted substantially from a briefing by Robert Lutz of Johns Hopkins University's Applied Physics Laboratory, a briefing given at the MORS minisymposium referred to in footnote 1.

TABLE E.1 Levels of Campaign Models

Level of Model	Scope	Level of Detail	Time Span	Outputs	Illustrative Uses	Examples
Theater/ Campaign	Joint and combined	Highly aggregated	Days to weeks	Campaign dynamics (e.g., force drawdowns and movement)	Evaluation of force structures, strategies, and balances; wargaming	CEM, TACWAR, Thunder, JICM
Mission/ Battle	Multiplatform, multitasking force package	Moderate aggregation, with some entities	Minutes to hours	Mission effectiveness (e.g., exchange ratios)	Evaluation of alternative force-employment concepts, forces, and systems; wargaming	Eagle, Suppressor, EADSIM, NSS
Engagement	One to a few friendly entities	Individual entities, some detailed subsystems	Seconds to minutes	System effectiveness (e.g., probability of kill)	Evaluation of alternative tactics and systems; training.	Janus, Brawler, ESAMS
Engineering	Single weapon systems and components	Detailed, down to piece parts, plus physics	Subseconds to seconds	Measures of system performance	Design and evaluation of systems and subsystems; test support	Many, throughout R&D centers

work and models are consistent with what he would obtain if he could do detailed analysis. So also, those who work at relatively high levels of detail are concerned about real-world contexts and constraints, which may be limiting factors in determining how systems are used and how they will perform (e.g., whether fighter aircraft will be permitted to engage at beyond visual range).

At any given level of activity, we need a model of how the world works that depends only on variables at that level of activity. For example, commanders maneuver forces and fires, and allocate other resources, defined doctrinally at their level. They must limit complexity if they are to operate effectively. They *care* deeply what goes on at higher levels of detail, but they can only check on such matters by exception. Instead, they must depend on doctrinal planning factors, aggregate models, and judgment with occasional high-resolution "calibrations."

It is worth noting here that most analyses and exercises depend on being able to treat key phenomena in relatively higher detail than other, less-central phenomena. For example, in one campaign analysis, logistics may be represented by nothing more than supply and use rates (both in tons per day), while combat forces may be represented at the level of brigades, squadrons, and missile ships. In a logistics-oriented campaign study, this situation might be inverted, with combat being represented by a simple demand function, and logistics represented in some detail by airlift, entity-level sealift and logistics ships, and intra-theater distribution systems.

Distinguishable Problems

Assuming interest in having and linking models of different resolution, there are a number of related but distinct problems. These include the following:

- Making selectable resolution feasible and sound within a distributed-simulation environment where there is need for repetitive aggregation and disaggregation.
- Making selectable resolution feasible and sound within an analytical model where certain subroutines need to be at higher resolution than would be appropriate generally.
- Developing sound mutually calibrated families of models so that, at each level, work reflects the full range of available knowledge.

For each, there is a distinction between working with existing models and designing new ones.

EXAMPLES OF HOW THE ISSUES ARISE

So far, our discussion has been abstract. Let us now provide more concrete examples.

Improving the Basis of Parameters Used in Higher-level Analysis

Suppose one is assessing the potential value of a force posture dependent on naval and Air Force aircraft and on long-range missiles with precision weapons (e.g., missiles that might be launched from an arsenal ship). An operational analysis for a JTF commander might use models with factors such as the average number of aircraft sorties per day and the average number of armored vehicles killed per sortie. By contrast, a high-resolution simulation might consider variables such as the weapon configuration on each type of aircraft, the distance they must fly from aircraft carriers or bases, the tactics of maneuver (including concentration in time and dispersal of vehicles), and the capabilities of reconnaissance and surveillance systems. Both levels of resolution (and others in between) are respectable and important. However, estimates of, say, kills per sortie should be based on something more than conventional wisdom and Service claims. Too often, there is no documented basis. Further, there is no integrated family of models that would provide such a documented basis. Such a family is needed because the gap between test-range data and campaign effectiveness is too great for the connection to be drawn easily.

How Multiple Resolutions Arise in Simulations

Multiple resolutions are needed even within individual simulations, especially in the distributed simulation environments central to the future of DOD's M&S. Some examples of why follow.

- *Different echelons.* Some ground-warfare component simulations represent individual platforms as distinct entities, while others represent higher echelons as distinct entities. For example, semiautomated-forces models may represent tanks, while a corps-level combat model represents either companies or battalions. When these components are connected in the distributed simulation exercise, problems arise when a platform object needs to interact with an aggregate object. They also arise when aircraft entities need to interact with aggregate ground combat entities, or with naval entities. This cross-service issue makes it a greater concern for JSIMS.
- *Different levels of detail of entities.* Even at a single echelon, entities may differ widely in the level of detail they represent. A basic aircraft simulation might represent only 3 degrees of freedom (DOF), such as X, Y, Z and their rates of change. A more detailed model might represent 6 DOF (X, Y, Z, yaw, pitch, roll, and their rates of change). Both support interactions with a simple range-only sensor model, but only the 6 DOF model supports a detailed sensor model that uses orientation to compute signature and detection probability.
- *Different processes within objects.* Even if a simulation can always represent interactions between entities at the same level of resolution, the desirability

of simulating those interactions at all may vary over time. For example, a logistical base may simply use integer counters to model the cycling of equipment through various stages of readiness most of the time—but when the base comes under attack, it becomes important to represent those items of equipment as individual entities to be sensed and attacked. A C^2 node in computer-generated forces may use simple decision logic when simulating noncritical parts of the battlefield, but use sophisticated decision logic when simulating critical parts—even though the same kinds of physical entities and physical interactions are supported everywhere.

Practical Problems Arising in Distributed Simulation

The example above involved analysis, but there are also many problems that arise in distributed simulation intended for training and exercising forces and their commanders. Some are down-to-earth in character, but troublesome to simulationists who must do the best they can to construct a synthetic theater of war. Some of those problems are as follows:[3]

- *Differing time steps.* Suppose a semiautomated-forces model (e.g., ModSAF) runs with approximately 1-second updates for each entity, but is interfaced with a tactical-level model (e.g., AWSIM) that runs with approximately 1-minute time steps. What does ModSAF see between AWSIM updates, and how does AWSIM handle short-lived combat interactions?
- *Templating subobjects.* When a battalion object encounters a collection of tank objects, where does the battalion place all its newly created vehicles as it deaggregates?
- *Duplication of C^2 processes.* Do we need to write one computer-generated-forces (CGF) command-control rule set for a simulation when it is running battalion-level objects, and a whole separate CGF/C^2 rule set when it is running entity-level objects? This would imply near-duplication of programming and knowledge-acquisition effort, multiplication of scenario setup effort, and exponentiation of VV&A effort.
- *Results correlation.* If a combat process can be simulated at both high and low resolution, how can we guarantee consistency between the two results—even when the two processes start with the same scenario?
- *Consistency between repeated deaggregations.* If one object changes resolution several times in a row, how can we ensure that the sequence of detailed views represents a coherence sequence? For example, as a battalion deaggregates, reaggregates, and deaggregates again, how can we make sure that the subordinate platoons (or even subordinate tanks) do not jump around in physically impossible

[3] These problem examples were suggested by Ben Wise.

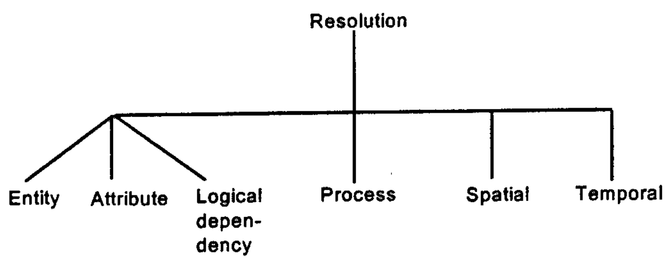

FIGURE E.1 Aspects of resolution.

ways? If the same logistical base is attacked several times in rapid succession, how can we ensure that the equipment on the base is properly placed? When do these issues matter?

- *Wide area sensors.* When one wide area sensor, such as JSTARS or overhead assets, views the battlefield, must everything in the whole theater change resolution to support that one sensor?

Against this background of challenges, let us now discuss what is involved in multi-resolution modeling.

FUNDAMENTAL ISSUES

What Is Resolution?

The difficulties in discussing variable resolution or multi-level resolution begin with the word "resolution," since resolution is multifaceted as Figure E.1 suggests. To make matters worse, in comparing two typical military models, one often discovers that the first model has higher resolution in some respects and lower resolution in others.

Usually, people doing simulation think of higher resolution as associated with lower-level objects (e.g., with individual tanks rather than aggregate concepts such as battalions). However, a "high-resolution" model representing individual vehicles might not distinguish among them, and it might assume they all moved in lockstep. Further, it might compute the attrition to vehicles by estimating a higher-level attrition (e.g., to battalions or even divisions) and then allocating that attrition among the vehicles. Such a model would have *low* resolution with respect to entity attributes, process, and so on. The point, then, is that "resolution" is a complex subject. This certainly applies to naval forces, because in some simulations a cruiser may be treated as a single object—in some respects analogous to a tank—whereas someone interested in the cruiser's armaments and sensors would see it as being a complex system with multiple levels of lower-level entities.

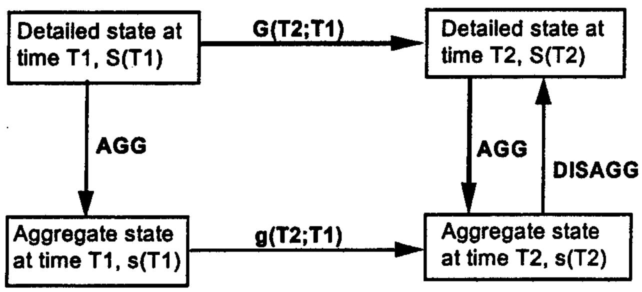

FIGURE E.2 Consistency diagram.

What Is "Consistency" in Multi-resolution Systems or Families?

A primary concept in MRM is that of "consistency." The recurring issue is whether two models—one of them having higher resolution than the other—are somehow "consistent." This is not a straightforward concept because the answer depends on context. Figure E.2 depicts the issues graphically. If G and g are high-resolution and low-resolution models, which operate on initial states to generate subsequent states, then the first question of consistency is whether one can start at the top left corner with an initial detailed state and get the same aggregate state by aggregating the initial state and applying the aggregate model (down and right) or by applying the detailed model and then aggregating (right and down). That is, we might hope that the aggregate model gets the same aggregate result as the more detailed model.

A tougher criterion for consistency would be requiring that the same final detailed state could be generated by moving down, right, and up, or by moving right. This form of consistency is more difficult to achieve because information is discarded in the aggregation process. How, then, does one regenerate detailed state information at the end? The answer, in some cases, is that the final state of the real system does not in fact depend on the initial detailed state. For example, if a carrier battle group moves from one location to another and then takes up battle positions, the spatial distribution of ships may be independent of the original detailed state, and dependent only on a combination of local information and doctrine—information added as needed. More generally, however, we have to expect that the second type of consistency will not be achieved.

Even here there are subtleties, however. Should the aggregate model really generate the same final aggregate state, or would doing so be merely accidental? After all, the aggregation of the detailed state is an aggregation of only one case, whereas the aggregate model may be dealing with averages over many cases. To be less abstract here, one would not really expect a detailed theater model to generate precisely the same overall attrition and movement as an aggregate model. Instead, one might expect that a statistical average over cases of the detailed

model's overall attrition and movement for a given case might be consistent with the predictions of an aggregate model.

The difficulties in formulating consistency illustrate issues that a theory of modeling and simulation should address (Appendix G). Conceptual clarity and mathematical rigor can be gained by applying such concepts as morphism and experimental frame, which such a theory provides. The basic concept of morphism, called *homomorphism*, is illustrated in Figure E.3. Two models are considered: S and S′, where S may be bigger than S′ in the sense of having more states. As in the consistency discussion above, when S′ goes through a state sequence such as *a,b,c,d*, then S should go through a corresponding state sequence *A,B,C,D*. We do not assume that states of S and S′ are identical—only that there is a predefined correspondence between them illustrated by the connecting lines in the figure. Now to establish that this correspondence is a *homomorphism* requires that whenever S′ makes a transition, such as from state *b* to state *c*, then S actually makes the *sequence of transitions* involving corresponding states *B* and *C*.

Some points to notice in this definition are as follows:

- The situation where S has many more states than S′ occurs in two major contexts: in multi-resolution modeling when S is a high-resolution model and S′ is a consistent (i.e., homomorphic) lower-resolution representation and in simulations where S is a simulation *program* and S′ is the underlying model.
- S may take a number of *microstate* transitions to make the *macrostate* transition from *B* to *C*. In the case of simulation, these are computation steps needed by the simulator to correctly execute the model state transition. In the multi-resolution case, both time and state are being aggregated in the lower-resolution model.

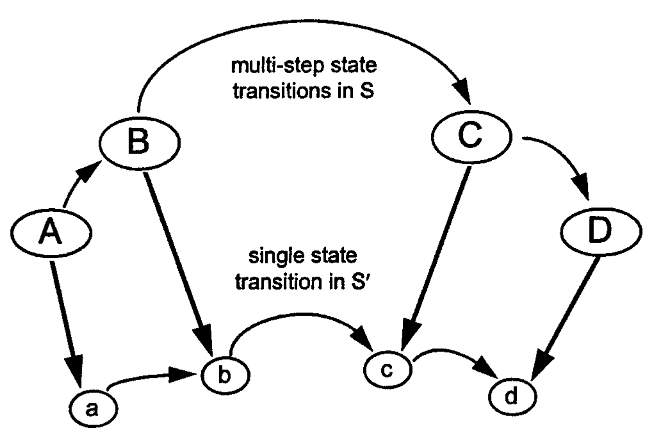

FIGURE E.3 State transitions in homomorphic models.

• Sometimes, we require strict step-by-step correspondence—i.e., that the transition from a to b is mirrored by a one-step transition from A to B. This is the case where both models are required to operate in strict time synchrony, as might be necessary in a real-time application.

• Typically, only a subset of the states in S correspond with those of S′. This subset is the *operating region* of the homomorphism. In the multi-resolution case, the operating region is the domain of the high-resolution model for which the low-resolution counterpart should be valid. For example, a high-resolution model of a fluid undergoing laminar flow may have a low-resolution representation, whereas its turbulent regimes may not. As discussed later in Appendix E, this is one place that the concept of *experimental frame* enters: an experimental frame specifies the operating region in which the low-resolution model must be a valid representation.

• Also, to achieve true abstraction, the correspondence between states must be many-to-one; that is, many states of S correspond to the same state in S′. For example, there may be many detailed states in the circle labeled B that are all represented by the same aggregated state b. In this case, the mapping from S to S′ does not have an *inverse.* In other words, as mentioned above, where true abstraction is involved, disaggregation is not a unique operation.

• A second important place where the concept of *experimental frame* helps bring clarity is in the relationship of the complete state of a model and its observable output. An experimental frame specifies the variables in which we are interested for some particular exercise. If a high-resolution model has the capability to compute such variables, it is indeed applicable to our frame of interest. However, the high-resolution model may do this in an "overkill" manner, and it may also compute a host of other variables that are not of interest in our frame. In this case, we may expect that a homomorphic low-resolution equivalent may exist.

Creating Integrated Model Families

Assuming we can define resolution and consistency in a context, a central challenge is developing integrated model families. How to develop these families is a frontier issue.

We may start by asking what we mean by "an integrated family of models." First, we mean that depictions at different levels of resolution are appropriately *consistent* or *morphic* in one of the senses discussed above. We also mean that data can flow meaningfully from one model to another, either by connecting the models as software or by having humans turn outputs from one into inputs of another. The word "meaningfully" is significant because, in practice, it is often not evident how models purported to have a family relationship should be connected. The models have often been designed with different perspectives on how the world works, as well as with different meanings for the same word or phrase

(e.g., "force ratio"). Or, it may be that the models were constructed with different operating regions of validity.

Yet another characteristic of integrated models would be that the variable names and function names would be conceived within the same global view, from top to bottom, thereby making it much easier to understand what a given variable means and how it relates to variables above and below it.

Note here that the goal of integration is not to create "seamlessness" (impossible), but rather—as suggested to us by John Doyle—to create "good seams," so that moving across levels of resolution maintains a clear and consistent sense of the system.

Integration of models has always been desirable, but analysts working in a single small organization have often been able to work around problems by studying the various models in detail and developing "good-enough" procedures. They have taken shortcuts and sometimes made errors, but at least the situation was to some extent under control. By contrast, consider the situation with distributed simulation. Here workers in different organizations are using data from each other's models and hoping that they are doing so sensibly, but without having full familiarity with all the pieces—and without even knowing the individuals who created the pieces. This makes the needs even greater than ever before.

HIGHLIGHTS OF PREVIOUS WORK

Having discussed some of the most fundamental issues, let us now review briefly some of the conclusions available from previous work. We highlight some that bear on common misunderstandings.

Misconceptions and Red Herrings

1. *Just building a good high-resolution model is not the answer, even with fast computers.* To many people, it seems as though the answer is simply to "do it right" with a high-resolution model and, as necessary, to generate aggregate displays. That, however, is wrongheaded. First, we do not have the knowledge necessary to build the requisitely comprehensive high-resolution wide-scope models (e.g., the knowledge to represent human behaviors well). Second, even if we did, we would not have the necessary data. Indeed, many of the critical data are unknowable in advance. Third, even if we somehow had the model and all the necessary data, we could often not do analysis without aggregating and smoothing.[4] And, to do that, we would need to know *how* to do the aggregation

[4]As one example here, the exploratory analysis emphasized in Appendix D is not feasible without abstraction (aggregation) because the curse of dimensionality is overwhelming even with massive computer power. With multi-resolution designs, however, exploration can first be accomplished with relatively abstract intermediate variables, and then refined by "zooming in" on those subordinate high-resolution variables of most importance.

and smoothing. Fourth, even if we could do all that, we would not know whether to believe the results or how to understand them, because the "explanation" would be at the level of bullets and trees. That is, we might have to construct aggregate models to comprehend and explain.

In summary, the problem here is not with computer speed, but with matters more fundamental.[5]

2. *Pure bottom-up approaches fail.* For related reasons, efforts to build complex system models strictly from bottom-up details have generally failed—collapsing under the weight of data requirements and shear complexity. Despite heroic efforts, they have often not been able to generate macroscopic behavior (recall Clausewitz's discussion of friction in war). By contrast, approaches that freely mix top-down and bottom-up approaches have a better track record (e.g., approaches that build in command and control structures from the top down). Further, recent work suggests that it is useful to think also about minimizing some details at the bottom of the bottom-up effort. Sometimes, it appears that only a few key features of entity-level behavior really matter to macroscopic behavior. The point here is that past experience, as well as theory, indicates that "purist approaches" based on strictly bottom-up (or, for that matter, strictly top-down) attitudes should be resisted. To represent complex systems well, one must use information from all levels, and welcome doing so rather than regarding some of it as the application of fudge factors. It is also important to be open to the need for iteration, because which entities make sense is sometimes not apparent until one has considerable experience, including experience observing so-called emergent phenomena.[6]

3. *Object-oriented programming will not solve the problem.* Object-oriented programming is excellent for describing hierarchies of natural objects (e.g., the carrier battle group that breaks down into component ships). However, the hard part of variable-resolution modeling or developing integrated families of models lies not in the object description, but in the description of how *processes*

[5]It is significant that physicists do not explain the skidding of an automobile in terms of Schrodinger's equation. They work with engineering-level equations and concepts such as the coefficient of friction, which they measure. Similarly, much of our best knowledge of military operations comes from aggregate-level observations and is expressed in the concepts of aggregate models. The commonly held notion that the best information resides only at high resolution is wrong.

[6]We base our comments here on our experience, our sense of the literature, and very helpful discussions with fellow panelist John Doyle and with Chris Barrett and Darryl Morgeson of Los Alamos National Laboratories (specifically about their experiences with the TRANSIM modeling effort to represent automobile traffic in large cities, with both detailed bottom-up modeling and a more agent-based approach using cellular automata). For an excellent semi-popular description of agent-based modeling and emergent phenomena by one of the pioneers in the study of complex-adaptive systems, see Holland (1995). For a recent survey of related work and its potential relevance to military problems, see Ilachinski (1996a,b). For a good entry to the important work on complexity of the Sante Fe Institute, see its Web page (www.santefe.edu).

at different levels interact. To use the example we started with, how does aggregate-level air-to-ground effectiveness relate to entity-level factors such as single-shot kill probabilities for precision weapons launched from 10-km altitude on a foggy day against a tank in the open? Relating these is analogous to relating thermodynamic relationships to the relationships of molecular physics and chemistry. Actually, it is harder, because in physics averaging in the process of aggregation does not have to contend with living, thinking, competitive warriors who are attempting to *avoid* things "averaging out" (e.g., by concentrating forces).

The point here is that we need humility in taking on the challenges of aggregation and disaggregation. Remarkably, modelers often display more hubris than humility in this regard. "Designing on the fly" at the computer terminal, they do violence to the underlying phenomena as they assume aggregate relationships that ignore complications and assume, implicitly or explicitly, circumstances such as uniform distributions, independent events, and constant remixing. Similarly, high-resolution modelers sometimes ignore frictional processes and give only short shrift to the all-important issues of higher-level command and control decisions. Whether one programs with an object-oriented language is irrelevant when the real difficulties are phenomenological.

On the Need for Hierarchical Designs

One possible solution to design challenges is called integrated hierarchical variable-resolution modeling (IHVR) (Davis, 1993). When feasible, it simplifies and clarifies the problems associated with crossing levels of resolution, either within a single model or within an integrated family. The basic idea is to design the models so that a given key high-level variable is expressed as a function of lower-level (higher resolution) variables, each of which is in turn a function of lower-level variables, recursively down to the lowest level. Ideally, this generates perfect hierarchical trees in which a given variable relates to variables above it and below it in the same tree or subtree, but never to variables in another tree or subtree. There is no cross-talk.

Figure E.4 illustrates the basic concept with a simplified representation of the ship-defense problem.[7] Suppose one is concerned about the probability that a particular ship (e.g., an Aegis cruiser) survives an attack by enemy ballistic or cruise missiles. In some war games, one might just specify that probability as a parameter, varying its value to see the consequences (Level 1 modeling). More

[7]The discussion here assumes, for simplicity only, that the incoming missiles can be treated independently. This is not true in practice, and a more serious treatment would require considering salvo tactics, saturation effects, and so on. The result would be a blurring of the levels and a blurring of the concept of leakage. As one possible outcome, a "correct" aggregation—or at least a good approximation—might involve a leakage that was a function of the number of attacking missiles and their "type tactics."

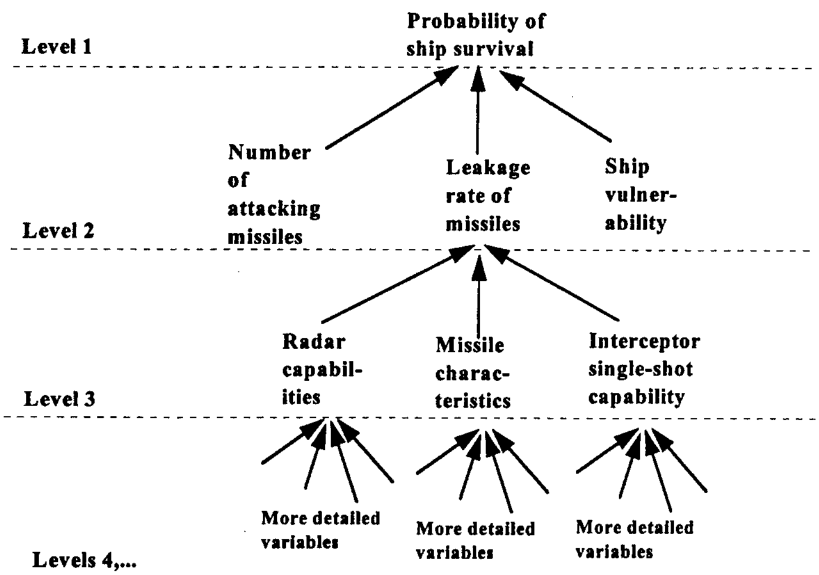

FIGURE E.4 An illustrative model hierarchy.

typically, one might have a model that calculates the probability of ship survival as a function of the number of attacking missiles, the leakage rate of those missiles, and the ship's vulnerability (i.e., the likelihood of being disabled as a function of the number of missiles that strike it). This would be Level 2 analysis, with leakage rate specified as a parameter, and perhaps varied. But leakage rate could be calculated from more detailed factors if the information were available. It could be calculated as a function of radar characteristics, missile characteristics, and the single-shot kill capability of its interceptors. And so on, down to more and more levels of detail. Now, the hope would be that the estimates of ship survival would be "consistent" regardless of how the calculation was made. This would be possible if the probability distribution of leakage rates assumed at Level 2 was generated from Level 3 analysis—averaged appropriately over all the relevant operational circumstances. In some cases, it might be adequate at Level 2 to use a "best estimate leakage" and an uncertainty range, without the embellishment of a probability distribution.

If one has this type of design, then it is easy in principle to proceed. One can run the model starting at any level of the tree, treating the lowest-level variables at that level as parameters. The values of these parameters should then be made consistent with an appropriate context-specific statistical average (or probability distribution) over results of running the model at the next lower level (higher resolution). This consistency should be obtained by adjusting models and data at all levels in the tree to represent "hard" information at whatever level of resolution it is found. For example, the reliability of a complex weapon system may be

based not on laboratory experiments, but on the experience of dozens of military units over time.

Another benefit of this approach is that it is straightforward to define and give names to variables without getting them confused. For example, there may be a half-dozen different force ratios in a ground combat model, but each would have the necessary adjective to distinguish it.

Unfortunately, there are three basic problems in trying to achieve this ideal of IHVR. First, aggregation and disaggregation are conceptually difficult and often quite subtle—not only in military modeling, but also more generally. Consider here the efforts that have gone into deriving respectable mathematical expressions for thermodynamic-level characteristics of nature from the molecular laws of physics and statistical mechanics. These problems have been considered hard even for equilibrium systems and a Mother Nature who is not trying to complicate things.

The second problem is that there typically are many complex interactions in a realistic simulation model, interactions that violate the image of pure and independent hierarchical trees. In military affairs, for example, one might think that one could treat Army, Navy, Air Force, and Marine forces as having their own hierarchical processes. However, an accurate depiction would show a good deal of cross-talk, even more as joint operations become the rule rather than the exception.

A third problem is that analysts commonly take different "perspectives" of the same problem depending on precisely what problem they are working. Attempts to impose a single perspective would make no sense. However, different perspectives imply different hierarchical depictions. For example, Navy and Marine officers often conceive command-control systems differently for air campaigns. So also, in some cases Marines might model their air forces as providing a kind of force multiplier rather than as destroying enemy vehicles at a certain rate per sortie. This might reflect a particular view of how the air forces would be employed (e.g., for suppression and as directed fire akin to artillery). That perspective would not "fit" well with Air Force models affecting ground combat, but it would arguably be just as valid. The conclusion here is that we should not assume that a given set of hierarchical relationships would always be "right." Analysts understand this viscerally, but simulation modelers sometimes tend to think of their preferred representations as being uniquely correct.

LOOKING AHEAD: NEXT STEPS IN UNDERSTANDING HOW TO DO VARIABLE RESOLUTION DESIGNS

Past militarily relevant work has contributed to a better understanding of the conditions under which various idealized aggregate models are or are not consistent with higher-resolution idealized depictions. Such work has only a limited potential, however, because it depends on toy problems such as problems de-

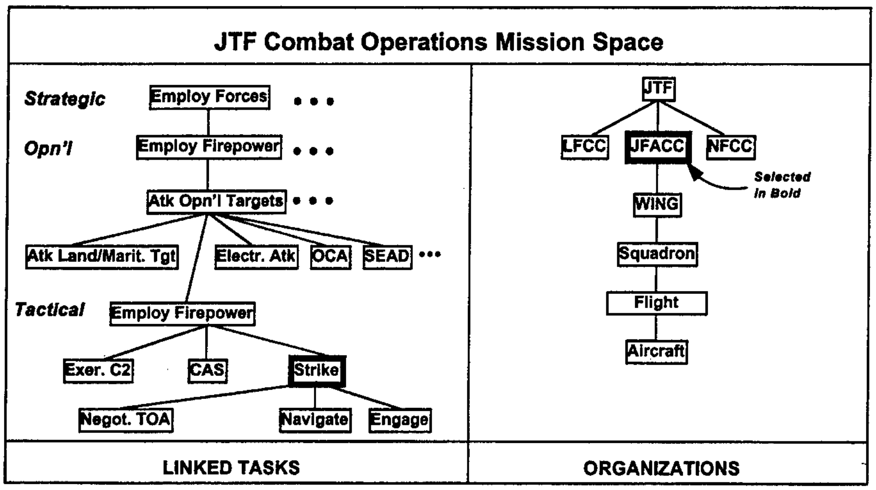

FIGURE E.5 A slice from a CMMS. SOURCE: Jefferson, DMSO (1996).

scribable precisely by Lanchester-square laws. What is most needed in the next phase of work is development of good *decompositions* and *approximations*. The world, after all, is to a large extent described by "partially decomposable hierarchical systems" (Simon, 1996). What are the right decompositions for military work? The answer is unclear, but the following are reasonable hypotheses.

The Value of Common Models of the Mission Space

Although workers in different aspects of DOD work will continue to have differing perspectives about how to characterize systems, there can be considerable convergence—to a small set of alternative structures rather than an unbounded set. Further, there can be agreement on the names to be used for commonly recognized entities and relationships, or at least on "standard" names and translations. Such developments would be quite valuable in efforts to build variable-resolution or multi-level resolution models.[8] Thus, the efforts of OSD's Defense Modeling and Simulation Office to encourage common models of the mission space (CMMS) should be supported. A good development here is agreement of the JWARS and JSIMS program offices to work on CMMS jointly. Figure E.5 (DMSO, 1996c), taken from DMSO materials, illustrates what is at issue. It is not exotic; rather, it is communicating concretely a vision of how the real world works. It shows breakdowns of tasks and organizations for a joint task force (JTF).

[8]Simple, worked-out examples illustrating these issues are given in Davis (1993).

Exploiting Relevant Temporal and Spatial Scales

A second hypothesis is that great strides will be made in MRM only by exploiting natural temporal and spatial scales, some of which need to be identified and defined. As noted above, real-world processes are often interconnected, making hierarchical modeling and MRM very difficult. However, if one breaks the simulation into appropriate temporal and spatial chunks, it is likely that simplifications can be made that will create *approximate* hierarchies. With luck and hard theoretical work, it may be possible to deal with the errors so created by making occasional adjustments in coefficients—much as is done currently as models adjust coefficients when forces maneuver from one type of terrain to another over a period of hours.

Finding the appropriate scales and ways to exploit them will not necessarily be easy because warfare operations have become quite complex as maneuver of forces has begun to give way to maneuver of fire, as lethality has increased, and as a relatively small number of C^4ISR systems have come to play an increasingly critical role. It is also plausible that aggregate models will sometimes not be as useful as in earlier days because the decisive events may be fewer in number and more highly correlated. None of this is clear, however, and in-depth research is badly needed (as discussed in Chapter 6).

Fortunately, it is sometimes possible for even a modest amount of theoretical work to shed light on confusing multi-resolution issues. As one example, a recent study used analytical expressions to show how the advantages gained from operational-level concentration depend on the relative time scales for C^4ISR, maneuver, and duration of battle (Davis, 1995). The work demonstrated that quite different aggregate-level laws would apply, depending on the relationship among time scales. Although the work used a highly simplified model assuming Lanchester equations, the basic principles demonstrated were more generally valid and the points made had not been well understood over the years.

Computational Experiments and Exploratory Modeling

Many insights can be gained by conducting simulations conceived as computational experiments. This is especially true when several groups approach the same problem, even with allegedly equivalent tools. These often produce surprises, even for experts. Further, they can guide development of better approximate models at lower levels of resolution (Hillestad, Owen, and Blumenthal, 1992; Hillestad and Juncosa, 1993).

With modern computer technology it is also possible to design huge sets of computational experiments in an effort to "explore" the space of possibilities and gain an appreciation for what matters and when, especially in the presence of large uncertainties.

"Solving Problems" by Avoiding Them

A different tack will often be critical in dealing with MRM issues. Rather than putting substantial effort into developing sound MRM relationships that can be used within simulations, it may sometimes be wise to adopt standards for distributed simulations designed to avoid the need to move back and forth among levels of aggregation. It may also be possible to design the entities of M&S to have a mix of high- and low-resolution attributes, with the entities "carrying along" just that subset of high-resolution information most needed for the interactions of the particular simulation. Ideas along this line have been proposed and pursued by both Paul Reynolds and his collaborators at the University of Virginia (Natrajan and Tuong, 1995) and Ben Wise of Science Applications International Corporation.

Flexibility

Modular design is essential and is facilitated by object-oriented methods. Given a sufficient library of modules, it may be possible to change representations (perspectives) from one application to another without too much special-purpose tailoring to adjust the relevant hierarchies. It seems unlikely that a hard-wired family of models will prove nearly as valuable as one that allows analysts with different problems to tailor the models suitably without great difficulty. Perhaps most of the alternative representations with real value can be conceived in advance, but it is doubtful. On the other hand, with appropriate configuration control and documentation, each well-conceived tailoring would produce a new option that others could use in the future. Thus, the broader notions of model modularity and repositories to facilitate reuse are also consistent with needs for MRM.

Primers

A basic problem in both the design and the use of model families is that most workers do not really understand what is involved in crossing levels of resolution. As examples of what workers need, and as an opportunity to reinforce points made earlier, consider the following. Often, workers seem to believe that all they need to improve results are some high-resolution subroutines to be called as needed in the course of running their more aggregate simulation. Suppose that such subroutines exist, however. They must be initialized with high-resolution input parameters, which are nonuniquely determined by the lower-resolution state variables. Which disaggregation should one use? Or should one instead run a large number of the high-resolution cases using different input parameter values, and then somehow average the results statistically? If so, what statistical approach would be suitable for the problem at hand?

More generally, how "should" one calibrate values of input parameters at different levels of a model family? The answer is not at all straightforward. As above, there are complex issues of statistical averaging, made more difficult by the fact that the humans in military operations are, as mentioned earlier, trying to avoid the circumstances in which everything averages out. Also, there are many different sources of information, some of it at high resolution (e.g., the number of weapons of a given type carried by a given aircraft on a given day's sorties) and some of it at low resolution (e.g., the typical frictionally caused delays in various command and control processes). How can all this information best be used? Aggregate-level data may have important implications for high-resolution models (e.g., the implication that unmodeled frictions slow processes up), and vice versa (e.g., a serious mismatch between the effective shooting ranges of the adversaries may mean that aggregate models based on Lanchester equations or anything remotely comparable will fail catastrophically under some circumstances, as happened in Desert Storm—in part due to poor practices by the Iraqi ground forces (Biddle, 1996)).

Currently, there are few relevant primers, especially in military work, but even in the community more generally. Such primers are needed.

Tools

Although the current problems are due more to intellectual shortcomings such as the lack of good theories than to technology, technology can also help a great deal. It seems very unlikely, for example, that workers will go about their calibrations without fast running models and appropriate tools to define cases and accomplish the relevant statistical manipulations.

State of the Art

Fortunately, the theory and tools supporting integrated families of models have been making steady but slow progress, although the advances are not well known to the majority of military simulationists. Early work in aggregation theory for economic systems dates back to the early sixties (Simon and Ando, 1971). Cale (1995) gives a recent survey of results of aggregation theory in the ecosystems simulation context. The theory states conditions under which error may or may not be expected as a result of aggregation. Since it is generically stated, it may apply to many military situations. Indeed, it bears some resemblance to the work discussed in Davis and Huber (1992) and Hillestad and Juncosa (1993). The underlying homomorphism mappings and hierarchical, modular construction techniques have provided the basic tools to construct families of models in both ecological (Zeigler, 1979a) and generic contexts (Zeigler, 1978, 1979b, 1993). Fishwick developed a software system to demonstrate the feasibility of a hierarchical multi-resolution approach to wire frame animation of human

body motion (Fishwick, 1986, 1989). A contemporaneous conference, "Enabling Technology for Simulation Science" organized by Alex Sisti of Rome Labs (www.rl.af.mil/Lab/IR/IRXtra/confpro.html) features a review of recent work on model abstraction and its latest developments.

CONCLUSIONS

The next generation of military models needs to be designed so as to produce integrated families that cross levels of resolution. This will require a good deal of theoretical effort involving mathematics, software engineering, and—perhaps most important—a deep understanding of the phenomenology coupled with an appreciation for how models of different resolution should and should not be used. Currently, the field lacks ability to apply the necessary theory, tools, and primers. However, there are insights in the literature that provide a foundation. What is needed is both further development of this and use of it in implementing actual simulation systems.

F

Model Repositories and Assembly and Integration of Models

Bernard Zeigler, University of Arizona
Paul K. Davis, RAND and the RAND Graduate School

BASIC CONCEPTS

It is a waste to have to reinvent the wheel each time a new car is designed. Yet as successive generations of simulations were developed in the past, such wasteful restarts from scratch were the rule rather than the exception. Nowadays, the advent of *object-oriented design and programming* has provided the technology to support *object repositories,* where objects may be *reused* time and time again. Models are stored in a database called a *model base.* Suppose that we undertake a project to construct a new model for given objectives. Then models that can serve as components for the new model are *retrieved* from the model base. Then to *synthesize* or *assemble* the new model, the components must be coupled together appropriately. When validated, verified, or otherwise properly accredited, the new model is *stored* in the model base so that it can be reused in the future (See Figure F.1 and Zeigler, 1990, for more details). Unfortunately, this scenario is easier to describe than to bring into common practice. Some of the issues that arise are as follows:

- How can a modeler discover models that are relevant to project objectives?
- How can models be designed so that they can not only serve their current purposes but also anticipate future needs?
- How can models be decomposed so that their components can be placed in the model base and recoupled later in different configurations (recall that models employed in dispersed geographic simulations can be distributed over computers in many locations, compounding the problem)?

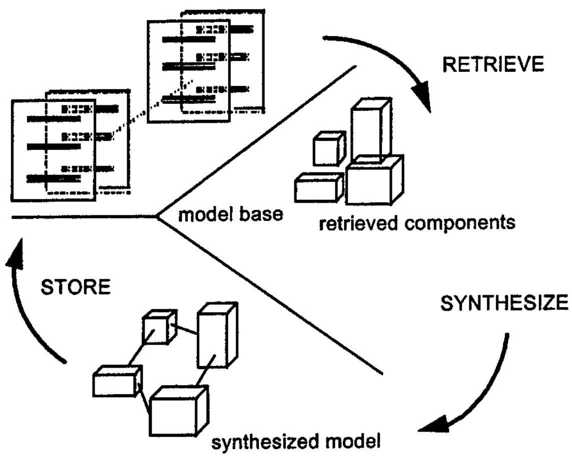

FIGURE F.1 Repository model base concept.

None of these problems is easily solved, but the modeling and simulation (M&S) framework provides some starting points:

- *Cataloging elements of the model base by type, application, and case.* Analysts and other users of M&S have long reused particular model versions and database versions. This is often referred to as using existing "scenarios," although that is an unfortunate use of the term scenario. However, the number of variations available, understood, and stored has typically been quite small (1 to 10, say, rather than hundreds). Further, it has typically been difficult to modify any of these stored models, in part because they have often been developed tediously so as to generate a particular "scripted behavior" involving large numbers of interacting entities and processes, which means that "small" changes can have repercussions throughout. In the future, much more should be possible.
- *Hierarchical modular model construction.* To be reusable, models must be self-contained with input-output ports as we have assumed in the system specification hierarchy. The model resulting from the coupling of its components must also be modular in this sense so that it too can be used as a component in larger models.
- *Building block components for application domains.* With some foresight it may be possible to design components from which a wide variety of models can be synthesized for a particular application domain. Thus, rather than focus entirely on the models needed for the particular project, model designers "regress" to a lower layer and search for good "primitives" to span the application domain.
- *Coupling templates.* Going hand-in-hand with the building blocks are standardized means to couple them together. The blocks must be designed to have the input and output ports that can be coupled together as assumed by the templates.

Reusability has obvious benefits in terms of millions of dollars potentially saved through faster project completions, and more reliable results with reduced manpower. Nevertheless, repository-based M&S has its costs in terms of specific design and maintenance requirements, as suggested above. Since these extra activities are not required for any particular project, they are likely to be considered a burdensome overhead for each such project. Given limited time and resources, a manager may be much more interested in completing the current project successfully than in laying the basis for the successful completion of future projects. However, an organization should adopt a long-term perspective in which the extra overhead incurred, especially in the first few projects, is traded off against the tremendous benefits that may accrue to future projects. In the context of advanced distributed simulation, multiple organizations may be involved in model development. The added complexity associated with coordinating individual efforts may greatly increase the difficulties in achieving reusability, while at the same time increasing the payoffs in doing so.

Models developed from systems concepts have identified input and output ports that enable them to be coupled together to form larger aggregates. However, models developed before object-oriented concepts took hold may be valuable, and it might be cost effective to reuse them as well. The hurdles in trying to salvage such *legacy* models (e.g., TACWAR and EADSIMS) are formidable. The problems in trying to interoperate or integrate a collection of such models arise from these complications:

- They may have been developed for disparate objectives, often not clearly stated.
- They may have made various assumptions, often undocumented, and possibly inconsistent.
- They may be built with varying levels of detail (resolution and scope).
- They may be implemented in disparate coded forms (languages, operating systems, and so on).
- Worse still, the experimental frame and simulation features may be tightly entangled with the model per se.

In contrast to the forward design of reusable object-oriented repositories, the backward retrofitting of legacy models may entail more cost than benefit. Sometimes it is possible to "wrap" a legacy model within an object interface so that it can properly interact with other objects. However, the prevalence of the above-mentioned problems may be so large as to make the effectiveness of such wrapping highly questionable. A more tractable integration may be possible where the outputs of models are not fed to inputs of other models but instead are employed to initialize their states or parameters. In this case, the models do not constitute components in a larger coupled model and do not have to meet the stronger requirements for consistent time advance and input-output compatibility.

DESIGNING FOR ASSEMBLY OF APPLICATION-SPECIFIC MODELS

In the discussion above we emphasized the synthesis or *assembly* of application-specific models from components. This may seem to be a straightforward suggestion, but it is distinctly at odds with traditional practice. Most existing large-scale DOD models of which we are aware were designed as a whole and are essentially monoliths. A few of the better-designed models have knobs and switches allowing some features to be turned off and on, allowing a run-time choice between high- or low-resolution depictions, but these are exceptions, and, even in these models, other complicated features are built in or interconnected in complex ways. The result has been that large and complex models have been used repeatedly for analysis that should logically have been done with much narrower models with fewer degrees of freedom. The old adage taught to all competent analysts is *that a model should be as simple as possible, but as complicated as necessary.* While the adage is widely given lip service, it is routine for it to be ignored by dyed-in-the-wool modelers and simulators, and even analysts who should know better, or who do know better but are stuck with monolithic tools.

Why is this so important? The answer is that good analysis depends on one or a very few minds completely comprehending what is being done. That in turn requires limiting complexity unless for some reason one can be confident that the various model components—and their data—are reliable. It would not be so bad if the large models' results depended on only a few uncertain variables, but the reality is that they may be sensitive to dozens, hundreds, or even thousands of uncertain data items of a large model. Some of the data for "peripheral aspects" of the problem may have been carefully established for different studies with different contexts, but may be quite wrong for the current study. But their inappropriateness may be difficult to uncover, and may insidiously corrupt the results.[1] Yet another reason for simplifying is that analysts must understand what they are assuming and what they are varying if they are to draw valid conclusions. Understanding the implications of large numbers of data assumptions is often impossible in practice. This seems unlikely to change unless model families are developed successfully.

For all these reasons and more, then, it is desirable for M&S to be designed for assembly. It can greatly improve reusability, quality, and controllability. Only a decade or so ago, it was extremely difficult to design for such features.

---

[1]As one example here, one might establish data values for many aspects of logistics if one were attempting to depict a best-estimate version of a particular war. In a subsequent study trading off alternative future forces and weapons, the outcomes might be strongly affected by the carryover data (e.g., one force might do poorly because it runs out of weapons or fuel, or is assumed to stop for a slow logistics tail) when the analysts are implicitly assuming that the future forces would be accompanied by suitable logistics. Such problems are common and insidious in monolithic systems.

That is no longer a limiting factor so long as maintenance can keep up with changes, such as those in operating systems and input-output programs.

Unfortunately for this story, the vision we are describing is much more suitable for high-quality (and highly paid analysts and M&Sers) than for "average" personnel, or even highly talented personnel with only short tours in a given position (a common problem for uniformed officers). Commercial desktop software may provide a familiar analogy. Desktop publishing software is highly flexible. People with desktop publishing skills can make almost anything happen, including changing page size, font, and orientation and importing graphics from many different authors and graphics programs. For most professionals, however, even highly educated and computer-literate "knowledge workers," there is value in having a stable, no-surprises software setup for text and viewgraphs, even if it lacks some desirable flexibilities. If models are used routinely for the same tasks, then their users will also want stability, but if they are often used to examine new methods or systems, or for diversity of purposes, modularity and assembly will be critical.

EXAMPLES FROM A 1980s-ERA SYSTEM

Many of the points made abstractly above can be illustrated in the history (both good and bad) of a major 1980s analytic war game, the RAND Strategy Assessment System (RSAS).[2]

The RSAS was a global analytic war gaming system. It could represent joint warfare in multiple theaters, even the "intercontinental theater" of global nuclear war. However, it was designed with the intention of serving many purposes and being as flexible as possible. Submodels were developed for air, land, and sea operations, as well as strategic mobility. These were building-block models. Other building blocks were decision models representing behavior of theater commanders and top-level military and political authorities.[3] The theater-commander models took the form of alternative adaptive war plans such as rigid defense at the inner-German border versus a defense strategy that permitted early fallbacks to the Weser-Lech "line" if necessary. Warsaw Pact strategies varied with respect to the sectors of concentration, the use of the Austrian corridor (a

[2]The RSAS no longer exists. After the disintegration of the Soviet Union, there was very little support for continued maintenance and upgrade. Further, the existing software became outdated as new operating systems and commercial graphics emerged. For these and other reasons, many features of the RSAS slipped into archives. However, a stripped-down and improved version of the warfighting models was developed and named the Joint Integrated Contingency Model (JICM). It is now being used, along with other legacy systems, for operational- and theater-level work by RAND, OSD (PA&E), the Air Staff, and the war colleges.

[3]These decision models amounted to "agent-based modeling" to use the current vernacular. Indeed, they were called Red and Blue agents because of the links to concepts in the artificial intelligence community.

high-risk, high-payoff strategy), and the use of airpower. Both sides' plans included nuclear options and adaptations to the other side's nuclear use.

Particular instantiations of the RSAS were created for particular theaters, notably Europe's Central Region and, to a lesser degree, Southwest Asia and Korea, and the "theater" of intercontinental nuclear war. These were constructed with relatively specific purposes in mind, for example, (1) evaluation of alternative force structures (e.g., to support analysis in support of the Conventional Forces in Europe negotiations), (2) characterization of the military balances, (3) evaluation of alternative strategies for theater- and global-level force employment, and, importantly, (4) support of joint war games at the various war colleges and National Defense University. These instantiations, once created, were then used repeatedly.

In any given application, however, there were many "coupling problems" to deal with. For example, the political-level models might choose to escalate as a function of the opponent's "level of conflict" on an escalation ladder. However, the analyst had to specify how the simulation would translate physical events such as the number, location, and time of nuclear detonations to "level of conflict." As another example, the two sides' theater-level decision models had to be given alternative adaptive war plans to choose among. Typically, some of these plans were built specifically for the given study. Each such plan and the decision rules for adapting or changing the plans typically involved some variables that had to be specified by the analyst (e.g., variables related to complex political judgments and associated military constraints). When the strategic mobility model was used, raw data on the capacity of various type aircraft for various type loads had to be translated into the terms used by the model. And, at the tactical level of combat, offline studies (or expert discussions) had to translate the complexities of sortie generation, C^4ISR, and weapon delivery into average kills per sortie for a type situation. The point here is that a great deal of the system was indeed reusable and modular, but a good deal of expert tailoring was almost always required for competent use. Precisely the same situation exists with theater-level combat models in extensive use throughout the DOD (e.g., CEM, TACWAR, Thunder, and JICM).

Significantly, while the developers of the various current models understood the desirability of reusability, it was not feasible technically to imagine large-scale reusability across research organizations. Instead, even with the relatively modern RSAS and JICM, transferability and confederation with other models are quite difficult because of peculiarities associated with, for example, representation of geography, the operating system, and many other factors. Technical problems such as multiple changes in Unix operating systems and the diminution of support after the collapse of the Soviet Union led to major features of the RSAS going into the archives.

In the future we can at least aspire toward much greater transferability and reuse because of the standards being created (e.g., the HLA). It is plausible and

even likely that object-oriented programming and modular designs consistent with the HLA will make it possible for future systems akin to the RSAS to have long useful lives. This, indeed, is what is hoped for in the JWARS effort. Whether that is achieved depends on the intensity of devotion to keeping the JWARS effort an "open architecture" that can readily accommodate alternative modules and, thus, evolve if newer and better representations emerge of important objects or processes. The panel's experience has consistently been that day-to-day and economic pressures are almost always in favor of relatively monolithic, not extremely modular, constructions. The reasons are apparent to anyone who has built computer programs with more concern about speed of completion, run-time speed, and "straightforwardness" than about expandability, reuse, modifiability, and so on. This has not changed. Another factor is DOD's frequent emphasis on agreed databases and configurational control, sometimes at the expense of quality. The Department of the Navy should establish a continuing policy of arguing for the modular assembly-oriented features of JWARS and JSIMS, and increase the emphasis on such matters in more Navy- and Marine-specific models like NSS.

CAUTIONS ABOUT CROSS-ORGANIZATIONAL M&S AND ONE-SYSTEM CONCEPTS

Despite the theoretical and practical strength of modern model-building concepts and technology, we note that it is an unproved hypothesis that such reusability will be meaningful and sufficiently low-risk to be used in distributed analysis. It would not be surprising if cross-organization model confederations used in distributed simulation 20 years from now were as untrustworthy and impenetrable as large monolithic models are today—when used for tradeoff analysis and other complex tasks. On the other hand, model confederations have already proved useful, for both training and analysis, in a variety of situations.[4] Generalizations are dangerous, and much depends on how DOD manages its M&S in the years ahead.

Another caution is that building-block approaches have their limitations. There are costs associated with having a system with too many choices, building blocks, and features. In principle, such a system may be able to serve many different masters, with each assembling the system they need, but in practice the system may be difficult to comprehend and ponderous—especially when attempting to serve applications across domains with different concepts, purposes, terminology, and measures of effectiveness (e.g., training, test-and-evaluation, and force planning). As a result, there will continue to be demands for specialized systems with only moderate flexibility. The one-system-serves-all concept should be viewed with considerable suspicion.

[4]Examples of successful use of confederations were given in a recent minisymposium (MORS, 1997). See, for example, the paper by Kent Pickett of the Army's TRADOC.

G

Components of a Theory of Modeling and Simulation

Bernard Zeigler, University of Arizona

The text of this report calls for further work in developing, extending, and communicating theories of modeling and simulation (M&S). This appendix sketches some key features that any theory of M&S should have. In particular, a theory should provide a basic foundation and framework, formalisms for defining and manipulating concepts, methodologies for representation and abstraction, and mechanisms for executing the models (e.g., turning them into computer programs). What follows focuses specifically on models of dynamic systems, that is, models whose variables change in value over time.

FOUNDATION

To deal with the foregoing issues, a theory of M&S needs to establish a mathematical, rigorous foundation upon which to base its formalization of the elements and relationships it has identified. Although the foundation will necessarily be more difficult to comprehend than ordinary language, its underlying concepts should be understandable to people who are not mathematical experts. The advantages of having such a rigorous foundation are readily stated. One concerns communication: many of the confusions that impede progress are due to terms, such as "model," that have different meanings across disciplines. A universally accepted theory of M&S would provide the common conceptual framework and vocabulary for people from different backgrounds to communicate effectively. A second advantage is that rigorous principles provide the means to tackle problems beyond the reach of more informal methods. The value of this is clear from other areas such as physics.

Some of the requirements that such a foundation should satisfy are as follows:

- The foundation should be general, and it should be expressive enough to subsume the great variety of special formalisms, languages, and modes of expression prevalent in M&S practice.
- The foundation should incorporate the concepts of dynamic systems theory. Dynamic systems theory has provided a uniform set of concepts that help to understand how objects change in time, that is, their dynamics, and how these behaviors are related to the objects' underlying mechanisms or structure. General systems theory represents the convergence of rich traditions, in areas such as control theory and automata theory, to a common mathematical conception of a dynamic system.[1]
- More specifically, models should be formulated as means to specify dynamic systems. That is, a model should be understood as a combination of equations, rules, and constraints that, when correctly interpreted, describes a unique dynamic system from the collection of all such objects.

FRAMEWORK

Any theory of M&S should establish a framework identifying and defining the key elements of M&S and their relationships. As indicated, the theory can employ the powerful foundation of dynamic systems theory to express these elements and their interrelations. In choosing what to identify as key elements, the theory should draw on the actual practice of M&S so as to highlight distinctions that are indeed significant. As examples here, it is important to distinguish among the real system, a model, a simulator (e.g., a simulation program or a hardware flight simulator), and what is sometimes called the experimental frame. The model is an attempt to describe aspects of the real system in a specific context such as estimating the likely time dependence of a real-system variable for any of a specified set of initial conditions. A simulation program might generate that estimated behavior using the model's equations, rules, and constraints. The experimental frame specifies the input stimuli, outputs of interest, and context of use. Thus, it is closely related to the concept of experimental design.

Any framework for M&S should facilitate discussion of meaningful relationships among key elements. For example, it is important to be able to discuss the validity of simulated model behavior with respect to the real system in a particular experimental frame. That is, validity is a relationship measured for a context. Another example of a meaningful relationship is whether a simulator such as a simulation program has been verified as representing the model adequately, again in the context specified by the experimental frame. Numerical approximations, for example, might be entirely acceptable in one frame, but a source of unacceptable error in another.

[1]For a review, see Pichler and Schwartzel (1992).

A full framework should identify just the right elements and relationships to facilitate all aspects of the practice of M&S—including aspects involving portability, reuse, and composability.

FORMALISMS

A framework should provide basic concepts, but theories must accomplish a good deal more—allowing workers to reason rigorously about issues, derive theorems, prove correctness of simulators, and so on. As a result, theories require formalisms. Formalisms are typically mathematical languages. One example is the predicate calculus.

Set theory is a common way to construct formalisms. Assuming use of set theory, a formalism for M&S should have a number of attributes:

- The theory should characterize the three basic types of simulation model (differential-equation, discrete-time (or time-stepped), and discrete-event (or event-based)) through use of set-theoretic formalisms, which should also expose their commonalities and differences.
- The theory should specify means of composing models in the basic formalisms from more elementary pieces. One means of composing models is by connecting outputs to inputs. Such coupling should work on well-specified input and output interfaces, without reference to internal structure.
- The basic formalisms should be closed under coupling. This means that coupling models expressed within a formalism should only produce composite models that can also be expressed in the formalism. This supports hierarchical construction, modular reuse, and hierarchical simulators.
- The theory should support combination and extension of the basic formalisms. An example of a combination is the formalism combining differential-equation and discrete-event formalism for hybrid modeling. An example of an extension is that in which a formalism allows models to change their structure over time.
- More generally, the theory should provide a methodology by which a new specialized formalism can be instituted by defining the subclass of systems that the formalism specifies.
- The theory should also provide a methodology by which the coupling of a new formalism is definable and its closure-under-coupling properties demonstrated.

SIMULATORS

In many cases, models will define relationships capturing key aspects of the system being treated. In themselves, however, they may not generate predictions. As an example here, Newton's laws do not themselves tell us how a falling

body's altitude will change with time. For that we need to compute the implications of the model.

Simulators are the computational devices (be they algorithms, programs, hardware, or networks) that execute models to generate their time behavior. A theory of M&S must deal with simulators:

- For each modeling formalism, the theory should provide a simulator concept that can execute any model in the class (i.e., generate the system's estimated time behavior). (One way to develop such concepts is through object-oriented frameworks.)
- Such frameworks should enable the development of verifiable, efficient, and interoperable implementations in an open-ended variety of contexts and platforms (e.g., parallel, distributed).
- The theory should provide a methodology by which the simulator framework for a new formalism is definable and logical correctness demonstrated. This means that there should be a way to define the simulator for a new class of models and provide its correctness for that class.
- The theory should characterize, and provide a means to estimate, the computational complexity of a model. Roughly, the computational complexity is measured by resources required by the most efficient simulator to simulate it.

REPRESENTATION AND ABSTRACTION

The theory should deal with the representation of systems as models and the abstraction of models into (usually simpler) models. There are mathematical concepts, called "morphisms" that provide the formal equivalent of the relations underlying representation and abstraction. For example, an isomorphism between two (mathematical) groups is a one-to-one correspondence between their elements that preserves their group operations. Such groups are said to be "isomorphic" or "equivalent."

- The theory should permit transforming a model expressed in one formalism into an equivalent expressed in a different formalism (e.g., differential equation models can be cast into discrete-event equivalents, which are computationally more efficient).
- Further, the theory should provide a methodology for characterizing the class of dynamic systems that can be represented by a formalism under a prescribed morphism relation (e.g., the class of piecewise constant input-output systems is known to be representable by discrete-event simulation).
- The theory should provide the basis for abstraction, that is, transforming a model into an equivalent with reduced complexity, within a specified experimental frame.

- The theory should provide an open-ended set of abstraction methods (e.g., aggregation and omission) and characterize their applicability. The theory should characterize, and provide a means to estimate, the simulational complexity of a model (typically, abstractions are intended to reduce such complexity).

ENCOMPASSING THEORIES

The theory should provide the elements of manipulation for more encompassing theories such as those of systems engineering, design, and management.[2]

[2]For additional reading, see Praehofer (1991), Zeigler (1976), Pichler and Schwartzel (1992) and Zeigler et al. (1993).

H

Areas of Research in Modeling and Simulation

Bernard Zeigler, University of Arizona

The text discusses research areas in three categories: (1) modeling theory, (2) modeling methodology, and (3) tools and environments. This appendix provides examples of research in each of these categories. The items shown are illustrative only, the point being to demonstrate something of the diversity of issues needing research.

MODELING THEORY

Simulation-based Design Research

Manufacturing control is traditionally approached with analytic/Markov methods for the creation of analytic models. However, using discrete event models to represent the machines, material handling, and input devices frees the modeler for experimentation with new and unique control methods. Users can make decisions by observing simulations using realistic "scenarios" of the manufacturing process and examine the implications of change (Zeigler, 1990; Cho and Zeigler, 1997). Because of the modularity of the approach, a wide variety of on-line control elements—including not only classic control mechanisms, but also neural networks, fuzzy logic, or expert systems—can be installed for performance analysis.

While model-based control is intuitive and can represent some of the deep knowledge employed of a human expert charged with directing a process, the approach of applying discrete event simulation and the requisite large-scale computing for automation is still in its infancy. Further research is needed to bring it to the point where it can support manufacturing styles such as flexible or agile

paradigms. DOD has many manufacturing processes and similar processes, such as logistics repair, that could significantly benefit from agile or flexible design based on discrete event simulation.

Dynamic Structure Modeling and Simulation

In many important physical and military systems, the system "structure" changes in the course of time. For example, biological systems such as growing plants, and social systems such as self-organizing organizations (one model for highly dispersed ground forces in the future), change structures over time. So also does a military organization that suffers attrition and reorganizes with a new command structure or a military organization that reorganizes and replans because of events making the original concept of operations obsolete.

Although significant research has been done on such simulations, current simulation languages do not support them. To represent such changes, they must be recast into parameter changes, and this leads to convoluted code that is difficult to verify and inefficient to run. Augmenting or replacing current simulation languages to support dynamic structure modeling would greatly increase the power of simulations to study complex structurally variable systems to gain true insight and predictability. This technology has been the subject of numerous investigations, but only recently has a first theoretical framework even been proposed and implemented. Thus, research that can contribute to a coherent usable methodology is at an early phase.[1]

Inductive Modeling

Inductive modeling attempts to infer a system's internal structure from data representing its behavior. Given that data collected from all kinds of systems are abundant, realizing a comprehensive inductive modeling methodology will be of significant importance to the M&S community at large. Within the military domain, it may be possible to generate rich databases from exercises and training activities mediated by distributed interactive simulations.

Despite a large body of research in inductive modeling, there is little agreement on any recognized inductive modeling paradigm. Several software implementations exist, including one developed based on a well-defined framework for inductive modeling, and implemented in a Artificial Intelligence Truth Maintenance system supporting nonmonotonic reasoning (Sarjoughian, 1995). This type of reasoning is needed to support flexible assertion and retraction of abstractions and assumptions in model building. However, this work has only tackled

[1] One example of work in this domain involves support to DOD's business reengineering, which must reflect the self-organizing formation of teams in business structures.

"toy problems," and it is imperative to apply it to some real application areas. Fundamental research effort is needed to bring about a useful and mature methodology to support a multitude of DOD present and future activities within the next couple of years

The present and future mission of the DOD provides real-world problems for applying and validating an inductive modeling framework. Potential applications span all of the M&S activities of interest to DOD with significant implications for model characterization from behavior and model abstraction techniques. Examples are Advanced Imagery Exploitation and Defense Automated Warning Systems, as well as many other areas requiring nonmonotonic reasoning about abstraction and assumptions. An inductive modeling technology would help DOD to address problems where conventional M&S is inadequate because of an abundance of data together with a lack of a well-developed scientific knowledge base and the M&S know-how to make sense of it.

MODELING METHODOLOGY

Experimental Frame Methodology

Experimental frames enable simulationists to translate the objectives and issues to be addressed into conditions under which a model or real system will be experimented with (Zeigler, 1976). As a major part of the initial requirements specification, experimental frames are critical to appropriate choices (e.g., level of resolution and accuracy) throughout the subsequent modeling and simulation effort. Experimental frames map into modules that actually do the experimentation (input generation, output summarization, and so on) when models/systems are operable.

While the concept of experimental frames has been around for some time, it is only recently that full support for their specification, manipulation, and management has been attempted. Experiment plans are supported in a Bomb Damage Assessment environment (Simard, 1996). However, such plans are formulated after model development, rather prior to it, as in true experimental frames. Some current environments support experimental frame construction as executable components but do not support the more abstract specification needed for symbolic manipulations.

DOD M&S efforts often are overly costly owing to their inability to make critical choices such as scope of representation and resolution level that should be driven by issues-oriented experimental frames specified in advance of model building. Moreover, archiving experimental frames and then matching them with existing models would enable a high level of model reuse.

Automatic Model Verification

Automatic model verification (AMV) differs from the conventional model verification methods in which verification is based on manually executed simulation runs. AMV aims toward automation of discrete event models verification. One promising approach is based on dual specification (Hong and Kim, 1996). The approach employs two specifications for a discrete event model: an operational specification for the behavior of a model and an assertional specification for its temporal properties. A model's verification is based on a language acceptance checking mechanism for which the assertional model constitutes a language grammar and the operational model acts as string generators.

Promising research in AMV has been performed. Although no software tool for AMV based on the dual specification approach has yet been developed, a prototype has successfully demonstrated the approach. Further research and development is needed to reduce the approach to usable tools.

Model Simplification Through Change in Formalism

Continuous systems are traditionally modeled with differential equation models. However, recent research has suggested that discrete event models may afford advantages for simulating continuous as well as hybrid systems (Zeigler, 1989). Several approaches exist for faithfully mapping differential equation systems into discrete event models such as analytic expression of transitions, application of algebraic solvers, and fuzzy representations.

A discrete event model, which meets certain steady state conditions, has been shown to be equivalent to a Markovian process. When analytic solutions are available for such processes, they can be solved in much less time than simulation requires. Markov lumped models can also replace their base model counterparts within the original simulation model, leading to more efficient simulation. Analytic expression of transitions has been shown to provide some 100 to 1,000 speedup over conventional time-stepped numerical integration (Moon, 1996). However, in many situations analytic (local) solution may not be possible. Therefore further research is needed to test general methods that do not rely on analytic solutions.

Simulations including both continuous and discrete event model components are common in DOD applications. For example, airplane motion is described with differential equations, while decisions of an intelligent autopilot are discrete. In such simulations, the speedups obtainable with a complete discrete event representation, with or without further Markov reduction, would enable simulations that are currently not feasible to be conducted. For example, it would be possible to simulate terrain models using digital elevation data from geographic information systems representing large areas in high enough resolution for realistic tests of sensor systems.

TOOLS AND ENVIRONMENTS

Environment for Simulation and Implementation of Discrete Event Control Systems

Discrete event system models have had a major impact on control system design for modern automation and real-time decision-making systems (Ho, 1989). The design of discrete event control systems usually employs discrete event simulation to verify functional requirements as well as to evaluate performance. Such simulation can be performed in discrete event simulation languages. Once simulation is done, the implementation of the designed discrete event system may proceed using a programming language, such as C or C++, which can be executed in real time. Since source code implementation totally differs from that of the simulation model, this approach to design cannot reuse the simulation model code in implementation. An ideal environment supports a close relation between simulation model and implementation code. In such an environment, a set of operating system-like system functions supports execution of a simulation model in real time. Thus, the same model analyzed in simulation can later be converted to real-time execution in a near-seamless manner.

Database Support for Simulation Model Reuse

Large-scale, complex-systems modeling often requires management of simulation models in an organized library or database (Zeigler, 1984, 1990). One major advantage is the potential for reuse of component models at different subsystem levels. Such model management can be effectively supported by employing object-oriented database technology. In this technology, a system can manage not only model structure in the form of coupling relations between component models, but also model behavior in the form of source codes or compiled codes. Such coupling relations and/or behavioral codes can be reused later on as building blocks to build larger models.

This technology area has already successfully been applied in the development of intelligent simulation environments. However, much research has to be done in order to apply the technology in the real world. For example, we need to develop a method for generating simulation models residing in an object-oriented database from modeling requirements and objectives.

Insertion of this technology would provide great benefits to DOD in large-scale, complex systems modeling, simulation, and analysis. It significantly reduces model development time by an efficient reuse of existing simulation models as building blocks.

SIMULATION-BASED OPTIMIZATION ON HIGH-PERFORMANCE PLATFORMS

Simulation-based optimization can be employed in most aspects of system modeling and design, as well as in higher-level decision-making processes. A wide variety of classic search and optimizing methods are available. In addition, there is now emerging a considerable literature on applications using nontraditional methods, which have both advantages and disadvantages. As examples here, evolutionary global optimization methods (Fogel, 1994), such as genetic algorithms (GAs) (Miachalewicz, 1992; Goldberg, 1992), were developed to apply the adaptive process of natural systems to search problems, and to develop artificial systems that mimic the adaptive mechanisms of natural systems. GAs encode a potential solution to a specific problem on a simple chromosome-like data structure and apply such operators as selection, recombination (or crossover), and mutation to the structure in the hopes of getting closer to the solution. Although regarded as merely "trendy" by some, GAs have been applied to a wide variety of search and optimization problems by many researchers. For example, a class of parallel GAs (Gorges-Schleuter, 1989; Pettey et al., 1987) for simulation-based optimization was applied to fuzzy system design, optical interconnection network design (Louri et al., 1995), parameter tuning, and model abstraction of a large-scale ecosystem model (Moon, 1996). However, system design problems typically require optimization of models having a large number of parameters, each requiring high precision. These parameters increase the complexity of the problem, and working with all the parameters at the same time often causes GAs (or any other optimization algorithms) to stagnate at local minima. Existing approaches cannot exploit information about performance impacts to search parameter subspaces in relation to their criticality. To address these problems, a multi-resolution search strategy in a distributed, high-performance simulation environment was developed (Kim and Zeigler, 1996).[2]

High-performance Parallel Discrete Event Simulation

Mapping large-scale discrete event models onto massively parallel architectures (Almasi and Gottlieb, 1989) requires the support of a higher level of abstraction in parallel simulation environments (Fujimoto, 1990). Recent approaches have employed object orientation to encapsulate the internode communication mechanism providing a user with a higher level of control (Zeigler et al., 1997). Mapping of models is also supported by its portability across platforms. Large-scale parallel and distributed discrete event simulation environments demonstrate the

[2]For further discussion of some of these issues, see also the last portion of Appendix B.

capability to address very complex and time-consuming simulation problems while providing a high-level interface. High-performance simulation environments have been tested on several models, including a spatial watershed and a large cluster of ATM switch models. The simulation can help analyze the complex interactions in models consisting of up to 10 million components (e.g., landscape cells or ATM switch elements). Speedups of the order of 200 times have been obtained so that simulations that require several days to run in conventional platforms can be completed in under an hour. There are numerous large simulations that could benefit from this technology, for example, air traffic control and multimedia communication design problems.

Distributed Simulation of Heterogeneous Models

Although distributed interactive simulation (DIS) protocols do not provide for strict global time preservation among federated models, the high-level architecture (DMSO, 1996c) includes a more controllable runtime interface. There are still many issues that must be dealt with in HLA (Morgeson, 1996). This motivates the development of a methodology for distributed simulation of models written in different simulation languages/environments that preserves strict time correspondence. Formalisms for discrete event models can be used as a common communication means. A software bus and an associated protocol based on such formalisms can provide an interface among legacy models in such languages as SIMSCRIPT, MODSIM, and SLAM. Proposed also are protocol converters, which support communication standards for such models. The methodology can be implemented using a network programming language such as JAVA. Insertion of this technology would provide great benefits to DOD in network-based distributed simulation of a large-scale system in which models of subsystems are developed in different languages/environments. It significantly reduces model development by reuse of existing heterogeneous models.

ADVANCED M&S ENVIRONMENTS FOR INTELLIGENT/ COGNITIVE SYSTEMS

Building models of intelligence, perception, and human performance has proved to be difficult due in part to the uncertainty in the psycho-physiological theories proposed to explain behavioral phenomena. Modern software engineering approaches such as spiral development suggest intelligent and cognitive model development using an incremental refinement approach (Young, 1992). They also provide the ability to develop multi-resolution models, although the underlying understanding of phenomenology is often the limiting factor. Recent developments in neuroscience have enabled us to envision behavior as the synergistic result of biological cells-neurons. Dynamic neural ensembles (DNEs) (Vahie

and Jouppi, 1996) provide a dynamic environment and the components necessary for the development of highly complex cognitive models aggregating cellular behavior to represent intelligence and learning.

DNEs are compositions of interconnected dynamic neurons. At a more abstract level, "holon" hierarchy models are being developed. Simulation environments supporting such models use object-oriented programming techniques to provide ease of parameter modification and specialization of both behavior and structure. Applications of DNEs to real-time learning, control, and decision making are currently being pursued. DOD systems and component designs for the 21st century will have to increasingly address the issue of human operability and performance. The development of autonomous systems capable of functioning in dynamic environments is also an issue of interest. The first issue, operability and performance, requires an approach that needs to be seamlessly integrated into design. The successive approximation provides a methodology for integration of cognition and intelligence into the systems design. New forms of neural and cognitive models, capable of dynamic behavioral modification, need to be explored to adequately capture flexible behavior.

Visualization and Significant-Event Detection in Discrete-Event Simulation

Any large-scale simulation is by definition complex owing to the size and diversity of the data. Events (in discrete-event simulations) represent a set of states (in one or more models) that are capable of influencing the states of other models in the environment. Therefore, an event may be determined as *significant* based on the values of specific state variables (in one or more models). Significant events are thus said to occur in a time period when a predefined set of conditions is met by a subset of the variables in the simulation. The user defines what he considers to be significant events using primitives and model parameters, before simulation. At run-time, event detectors sift through the data looking for significant events. This enables the user/model developer to effectively pursue his goal (conceptual or analytical). In essence, significant event detection allows any large-scale simulation to be viewed at various levels of abstraction, where the level of abstraction is determined by the significance of the event.

Due to the size and/or complexity of most DOD simulations, this technology would impact virtually all application areas where M&S is used. Being generic in nature, the concept could be modularized as an independent entity in diverse discrete event simulations. In battle simulations where planning, resource and personnel deployment, and communication are independent entities, there are too many data to track. The same model can be used by commanders in charge of each of the battle spaces where a significant event for one may or may not be a significant event for another, radically reducing their output data set.

Graphical Description of Discrete Event Model Behavior

Many good graphical tools are in place for discrete event systems modeling. Such tools use icons to represent predefined models, most of which support users to add a new model definition and an associated icon to the existing library. However, little has been done in graphical notation for behavioral description of discrete event models. An excellent example for such notation in discrete event modeling is a stochastic Petri Nets graph. In spite of its generality in modeling stochastic systems, Petri Nets is limited to modeling a certain class of discrete event systems. Thus, graphical notation based on a sound semantics, which is easy to use and understand, needs to be developed for the rapid and accurate modeling of discrete event systems. The graphical notation should include such information as state transition function, output function, and sojourn time function for a basic component of a discrete event process. Of course, the graphical notation should generate executable simulation codes.

Anytime/Anyplace Concurrent, Collaborative Support of M&S Life Cycle

DOD decision makers are faced with the challenge of declining budgets for manpower and material, and for demands for flexible, cost-effective operations to meet the challenges of the post-Cold War world. M&S is being applied not only at technical and engineering levels to meet such challenges, but also at higher levels such as work-flow automation and business reengineering, where many stakeholders are affected. To undertake effective M&S throughout its life cycle requires the active involvement of the various groups involved with model development, simulation analysis, and implementation. Unfortunately, tools and methodologies currently available from commercial vendors and consultants are primarily single-user tools that provide inadequate support for the collaborative team-based environment that characterizes modern organizations. Moreover, this support is virtually nonexistent for distributed work involving groups that are geographically dispersed.

Group support systems research has developed a network-based set of flexible software tools that incorporate basic problem-solving techniques such as brainstorming, idea organization, voting, issue analyzing, policy formation, prioritizing, and stakeholder identification. Electronic communications allow all group members, whether distributed or co-located, to make contributions to the group's task both simultaneously and asynchronously. Such technology increases organizational productivity by decreasing manpower requirements and cycle times in projects. The scope of projects can also be expanded to include participants from several hierarchical levels, thus improving organizational communication while facilitating approval for decisions. In a competitive environment

where success is dependent on teams working together, collaborative software will increase the productivity and effectiveness of these teams.

Research is needed to extend advanced M&S capabilities by embedding them in the distributed group support tools environments, to enable distributed groups to construct, analyze, and implement model-based designs in concurrent engineering fashion.

I

Combat Modeling Issues

Paul K. Davis, RAND and the RAND Graduate School
Donald Blumenthal, Gualala, California
Donald Gaver, Naval Postgraduate School

INTRODUCTION

The design of JWARS and other new combat models should raise numerous issues about modeling approach and phenomenology. To some extent this has happened, particularly with DOD's recognition that such next-generation models must represent the effects of the C^4ISR systems on which much modern defense planning is focusing. In many other respects, however, discussions to date have not converged and have too often been conducted at the level of "labels" used as litmus tests. Some of the labels dividing people in discussion include Lanchester models, attrition models, deterministic models, and configural theory. There have been numerous heated discussions on such matters because of the "Grand Canyon" that separates the domains of modelers and analysts working at different levels of resolution and, typically, on different types of problems. In this appendix we try to shed some light on the issues. Readers should understand, however, that there are chronic controversies on these matters, and no two authors are likely to emphasize the same issues. Although our examples pertain mostly to ground combat, the principles involved apply also to naval and air warfare.

MAJOR OBSERVATIONS

Lanchester Equations as Red Herrings

Despite the hundreds of papers written about them, Lanchester equations (as most people understand this term) are largely irrelevant to today's combat modeling by DOD, which uses computer simulations, not simplistic constant-coeffi-

cient differential equations such as the Lanchester-square-law.[1] Lanchester equations will probably remain quite useful for making particular points in the classroom (e.g., illustrating the power of concentration or the value of "crossing the T" in classic naval engagements) or theoretical papers, but to argue about their more general validity is to chase red herrings. It is the simulations, not the Lanchester differential equations, that should be examined.

Today's higher-level combat simulations (e.g., those at division, corps, and theater levels) are best seen as implementing aggregate state-space models (something much broader than Lanchester models). The basic notion is that the "state" of the system (the two opposed forces, their strategies, and the environment in which they fight) can be represented by a collection of variables such as counts of personnel and vehicles in an area, and terrain factors characterizing that area, rather than the locations and current behaviors of all the individual entities such as individual soldiers and tanks. Usually, the simulation then generates the predicted future state as a function of the current (aggregate) state. In more general formulations, there can be "memory effects" of previous states as well. Again, the variables affecting this prediction are not just the sides' strengths (much less their scalar strengths, as in the simpler Lanchester equations). Instead, the predicted change of state depends on many other factors such as terrain, defender preparations, flank exposure, strategy, and tactics. One important change of state, typically made at the end of time periods or when some significant event occurs, is a change of strategy or tactics (e.g., a decision to attack or withdraw, or to maneuver reinforcements to a trouble area). It is then true that the close-combat ground-force attrition in a given time step is sometimes approximated by a local use of some Lanchester equation, but the "coefficients" used can be highly situation dependent, that is, dependent on many other state variables that change over time (Allen, 1992, 1995). Thus, the simulation does not (or at least is not intended to) behave like a constant-coefficient Lanchester equation.[2]

Breakdown of Aggregate State-Space Models

It has long been a reasonable hypothesis—but only that—that a relatively aggregated close battle in a particular area will have attrition that can be reason-

[1]The principal reference for discussion of Lanchester equations is Taylor (1983b), which also covers many generalizations of the original work (Lanchester, 1916), including generalizations such as Bonder-Farrell theory (Bonder and Farrell, 1970) used in simulations. See also the recent collection of papers in Bracken et al. (1995), which includes historical analysis, a translation by Helmbold and Rehm of work by Osipov, and considerable thoughtful discussion. Wise (1991) explains some of the fundamental ambiguities in using and calibrating Lanchester laws. Hughes (1986) and Deitchman (1962) discuss applications of Lanchester models to naval and guerrilla warfare, respectively. Dupuy (1987) includes discussion of how his extensive history-based work on combat modeling relates (and does not relate) to Lanchester theory.

[2]In fact, simulations do sometimes generate behaviors that look remarkably like what could be generated by such an equation, but that is an artifact of the particular application.

ably approximated by a state-space equation, that is, an equation relating the change in the sides' strengths (attrition) to various state variables and the duration of the time step, using the initial state-variable values of the time step and treating the combatants as all "in" the same location. The hypothesis clearly breaks down at low level (e.g., when evaluating alternative weapon systems in engagement-level combat where configural effects can be dominant (see also Appendix J)). The validity of the hypothesis also depends on there being many discrete countervailing microscopic processes (concentration and counterconcentration, ambush and withdrawal, fire and counterfire, and so on) that, over the time step and over the many replicas of the close battle across a theater, average to something relatively simple. This aggregate result may or may not correspond to a Lanchester square law, linear law, or something similar. It may be better described by the more general Bonder-Farrell equations, for example, but in some instances, there will be no such simplification because one side or the other has an asymmetric advantage that can be exploited because of multimodal probabilistic effects.

The issue, then, becomes where and when various aggregate state-space models provide a good approximation of aggregate-level phenomena. It is inappropriate to draw broad conclusions, because contextual details matter a great deal to whether and which aggregations make sense.

Myopia Caused by Head-on-Head Attrition

While accusing simulation models of being Lanchesterian is often misleading, what critics who refer derisively to Lanchester models actually have in mind (clearly or dimly) is often something else, that most of today's theater-level models were designed from a so-called attrition perspective that conveys an image of war as mere head-on-head ground-force encounters with the two sides fighting to the bitter end. That is in contrast with a maneuver perspective in which campaigns consist of the sides maneuvering their forces in an attempt to create favorable circumstances of battle and to extricate themselves from unfavorable circumstances. Sometimes, a maneuver strategy can achieve victory without an extended attrition battle because one of the sides finds itself hopelessly outpositioned—and perhaps weakened by loss of critical assets or a collapse of command and control and unit coherence (possible objectives of information warfare).[3] Unfortunately, strategy and maneuver are often simplistic in models and studies conducted with the perspective of head-on-head attrition. Skilled users of even old-fashioned piston models can represent many effects of operational maneuver, but in practice, the result is often less impressive.[4]

---

[3]For discussion of information-warfare effects on theater combat, see Bonder et al. (1994).

[4]For historical-empirical discussion of why standard attrition-warfare models are inadequate, see Rowland et al. (1996). Application to Korean analysis is described in Bennett (1995).

Theater- and Operational-Level Models Emphasizing Maneuver Warfare

Although the head-on-head attrition modeling perspective is common, some theater-level models over the years have been designed to represent maneuver explicitly. For example, the IDAHEX model (Olsen, 1976) used in the 1980s introduced hexes and reintroduced interactive gaming with human players to make operational-level decisions; it did so specifically for the purpose of focusing effort on maneuver.[5] The Army War Colleges used a simpler but roughly comparable model in the early 1980s for similar reasons (the MTM).

In the 1980s the RAND Strategy Assessment System (RSAS) was designed to focus attention on the strategy variable, introducing it explicitly in analytical war plans that included contingent branches and other adaptations. The RSAS also facilitated examining the consequences of nonattrition factors such as operational surprise, strategic flanking operations (e.g., Soviet use of the Austrian corridor), qualitative shortcomings in the fighting performance of some forces, the dependence of reserve-force effectiveness on training time before force employment (and the type of employment required of them), and the likely slowing effects of interdiction attacks. One version of the RSAS included a network model to improve the representation of flanking attacks, noncontiguous axes of advance, and critical nodes.[6] An improved version of the network model is incorporated in the JICM model, which has been used for extensive study of warfare in Korea, including warfare involving counteroffensives, flanking attacks, and asymmetric strategies involving weapons of mass destruction.[7]

Another maneuver-oriented model was RAND's TLC/NLC, which was developed to a prototype stage using object-oriented programming and advanced graphics (Hillestad and Moore, 1996). Among other features, it included a rich network structure and reflected the Soviet correlation-of-force methodology for planning operational maneuver.

While none of these models has been fully successful, while all of them share the severe shortcomings discussed below, and while even these maneuver-oriented models have sometimes been used in ways that reduce war to something looking like simplistic attrition warfare, the existence of the models and some of

---

[5]It is of interest to note that the developer, Paul Olsen of the Institute of Defense Analyses, was criticized at the time (1986) because IDAHEX was not a "closed" model and, therefore, was allegedly inappropriate for analysis. His view was that without representing maneuver, the various more popular closed models were inappropriate. In fact, IDAHEX was later used extensively for analysis by the SHAPE Technical Center and a few other organizations.

[6]For discussion of the RSAS, see Davis and Howe (1990), Bennett et al. (1992), and references therein. The network representation was due to earlier work by Patrick Allen and Barry Wilson. Some subsequent but unpublished documentation on RSAS 5.2 is also available through Bruce Bennett or Daniel Fox of RAND.

[7]The JICM model is documented in Bennett (1994) and subsequent unpublished materials. It is used at RAND, the Army War College, OSD's PA&E, and some other organizations.

the studies accomplished with them demonstrates that the state of the art in combat modeling is substantially more advanced than those who decry head-on-head attrition modeling and Lanchester equations sometimes suggest. It is also significant to note that even older piston-style models have been used creatively and realistically—not only in research studies, but also by the operational commands, including the U.S. Central Command when preparing for Desert Storm.[8]

Severe Limitations of Current Theater and Operational-Level Models

While many criticisms of current models are exaggerated or overgeneralized, there is consensus throughout the community that DOD's current theater- and operational-level models are severely flawed. The major problems include their being overaggregated; having primitive or no representation of C^4ISR, command-and-control, and information warfare; being almost exclusively deterministic; having too little representation of operational concepts, plans, and command; and having little ability to characterize the fluid and highly nonlinear combat operations anticipated for the future. Even many of the advanced features described above in connection with the RSAS and TLC efforts (notably those associated with decision models) no longer exist in operating models. Indeed, much of the current work with higher-level models such as TACWAR depends unreasonably on scripted representations of force employment, which are very difficult to work with because of the need for repeated iterations and tuning, and the absence of sufficiently adaptive behaviors.[9]

WHERE NEXT?

Research Opportunities for Improving Higher-Level Models

Ideally, aggregate models should be informed by and even derived from more microscopic theory and experiment, including simulation "experiments" conducted at high resolution. Such experiments have their own shortcomings, but can nonetheless be a rich source of insight.[10] Furthermore, they are now

[8]See articles by J.A. Appleget and F.T. Case et al., in Bracken et al. (1995).

[9]Users are quite aware of these problems, of course. In the recent Deep-Attack Weapons-Mix Study (DAWMS), the Institute for Defense Analyses used a linear program (WORMS) to ensure that allocations of deep-attack weapons in TACWAR would be in some sense "optimal." Other aspects of the simulated campaign, however, were much less adaptive.

[10]It does not follow that aggregate models *must* be derived from high-resolution models. Nor does it follow that some aggregate expression such as a Bonder-Farrel model or a Lanchester-square model, used locally, is invalid because some of those viewing the equations fail to see features they know are important microscopically (e.g., stochastic features, configural effects, and so on). More-

feasible as the result of advances in computer science and entity-level simulation. As discussed in the text, exploiting this opportunity should be given high priority.

While there will continue to be an important role for diverse state-space models implemented as simulations—including some that will continue to be incorrectly characterized as "Lanchester models"—a great deal of effort is needed to establish a better foundation for the assumptions used in those models. For example, many of the terrain factors used in combat models were estimated many years ago when it was computationally impossible to conduct high-resolution high-quality simulations in the numbers needed to identify good aggregate representations. So also, the assumption of deterministic aggregate behavior was made in part because it was computationally infeasible to do otherwise. That multimodal distributions aggregate into something simpler, which can be treated by deterministic equations (plus uncertainty analysis to account for important branches) was and is a reasonable hypothesis for higher-level battle, but we do not currently know when the hypothesis is correct. It is clear that it fails for engagements in which one side can consistently exploit a range advantage. In any case, a new round of such research is now possible and needed. While current DOD simulations are in some cases based on earlier research that included comparisons with higher-resolution simulations (see, e.g., Farrell (1989), which discusses the early development of Vector models), that work should be reopened since the quality of the high-resolution simulations is now so much better.

In pursuing a research program to connect the worlds of high- and low-resolution modeling, it is essential to recognize information resides at all levels of aggregation and to avoid a pure bottom-up approach. Instead, the ideal is an approach in which models of differing resolution are used to exploit all the information available and, then, to cross-calibrate each other. Constructing such mutually calibrated families of models is a major undertaking, as discussed in Chapter 6 and Appendix E.

In conclusion, while we have attempted to illuminate issues regarding Lanchester models and DOD simulations, and to soften exaggerations, we emphasize the need for in-depth research to better understand the phenomenology of combat. One important spinoff of this could be DOD's emerging with integrated or semi-integrated families of mutually calibrated models appropriate for the full

---

over, the usual claim that the validity of Lanchester equations depends on assumptions of homogeneous static forces with perfect local command and control is fallacious—the result of the classic blunder of confusing sufficient and necessary conditions. Lanchester equations are often motivated by simplistic models of combat, but the equations may be valid as aggregate-level descriptions in combat circumstances having none of the simplistic features said to be assumed. An analog here is that the ideal gas law does not depend on there being elastic billiard balls flying back and forth horizontally between walls (the "model" often used to motivate the law in high school chemistry). On the other hand, such aggregate depictions are clearly *not* valid in other cases and it is not currently clear when they are or are not.

range of M&S applications. The traditional approach of developing models separately for the various levels of resolution is fatally flawed when it means working with blinders on, which it often does. Those working exclusively at low resolutions are unlikely to understand the underlying phenomena and are therefore likely to misrepresent the aggregate phenomena. Those working exclusively at high resolution are unlikely to understand larger contexts and interrelationships. Their insights and conclusions may be much more conditionally valid than they realize. Further, the calibration of high-resolution models should exploit all the relevant information available, much of which is at low resolution.

We conclude, then, that research should be conducted jointly at multiple levels of resolution with a great deal of interaction and the goal of integration.[11] Such work should include relating stochastic and deterministic representations, as well as consideration of many other types of uncertainty. Such things will not occur without changes in both funding and management practices. The Department of the Navy should advocate such changes strenuously in the joint arena and with OSD. Otherwise, it is likely that the next generation of aggregate combat models for use in joint analysis will not be significantly better than the ones already available—and perhaps worse.

[11]This integration need not be in any single model, however. It might be in textbooks plus occasional cross-calibrations. We are not advocating single do-it-all-comprehensively models.

J

Probabilistic Dependencies in Combat Models

Paul K. Davis, RAND and the RAND Graduate School

BACKGROUND

One special request made to the panel preparing this report was for comments on the importance of configural models for mine warfare and, perhaps, for other classes of combat. The request reflected a decades-long controversy on how to represent mine and countermine warfare issues mathematically so that the models would be of value to the acquisition, training, and operations communities. The term configural model is associated with a series of studies using complex analytical probabilistic models, which account for many of the probabilistic dependencies that appear in mine warfare problems (Horrigan, 1991).[1] In a number of cases with real-world importance, the associated effects are quite large and cannot be accommodated by simple adjustment of planning factors. It is puzzling that there should still be controversy about the need to account for the effects explicitly. Discussion of the mine warfare issues is included in a separate panel report on undersea warfare (Volume 7 of the nine-volume series, *Technology for the United States Navy and Marine Corps, 2000-2035: Becoming a 21st-Century Force*), but in this report it seemed useful to point out that the issues are of a class that can be found throughout DOD. The need to account for such

[1]See, e.g., Horrigan (1991). Horrigan defines configural theory as "a mathematical theory for quantifying the relationships between the behavior of weapons in use in combat and their individual characteristics. Its principal purpose is to provide concepts and mathematical relationships to improve our understanding both of weapon behavior in combat and of combat effectiveness. Its name is derived from its central concept, configuration, which is the mathematical expression of the fact that the disposition in space and time of the targets and weapons of the attacker and the defender is inseparable from the outcome of the engagement and the combat effectiveness of those weapons."

dependencies or correlations is well understood scientifically and mathematically,[2] but many DOD models do not do so adequately. In what follows we give examples of such effects from other domains, primarily to illustrate their generic character and to thereby increase acceptance of the need to address them.

CLASSIC EXAMPLES OF CORRELATED EFFECTS IN COMBAT MODELS

The Fratricide Problem in Strategic Nuclear Warfare

In the mid-1970s, an important military issue was whether emerging Soviet ICBMs would be able to destroy U.S. Minuteman ICBMs in their silos. The answer depended heavily on the effects of targeting a given silo with two (or more) reentry vehicles, since neither accuracies nor reliabilities were high enough to assure high probabilities of kill with a single RV.

The naive calculation was to assess the probability of a silo's destruction D by n RVs as follows:

$$D = 1 - (1 - RP_k)^n,$$

where R is the reliability of a single RV and P_k is the single-shot kill probability for a reliable RV attacking a given silo (a function of the RV's accuracy and yield, and the silo's characteristics). The equation treats the RVs as independent. The second term is the probability that n independent RVs fail to destroy the silo.

The first problem with the naive calculation is that the reliability of a given RV is correlated with the reliability of its sister RVs on a given ICBM: the principal failure mode was not in fact the RV, but the missile. Thus, if a missile failed, all of its RVs would fail. As a result, it was often assumed that a nuclear attack plan would "cross-target" weapons so that a given silo would be attacked by RVs coming from different missiles. In that case, for two RVs the equation would be

$$D = 1 - (1 - RP_k)(1 - RP_k).$$

In reality, the problem is much more complicated because the effects of the successive RVs are *not* independent: the first RV, if it arrives and detonates, may create shock waves and send dirt and other debris into the air through which the second RV must penetrate. On the other hand, partial damage from the first RV may reduce the strength of the silo to a second, and so on. On a larger scale,

[2]In statistical mechanics, the term "correlations" is often used to mean what we refer to here as "probabilistic dependencies." In some fields, however, "correlations" refer to only a subset of the many possible dependencies. Thus, we have avoided the term here.

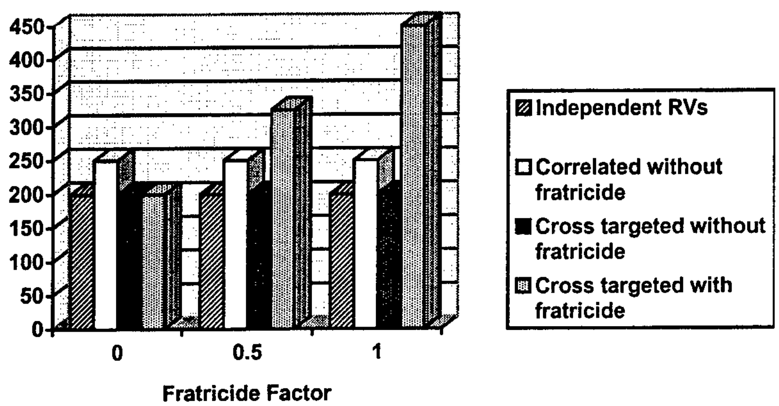

FIGURE J.1 Effects of fratricide.

detonations at one silo could affect the environment seen by newly arriving RVs at relatively distant silos. If such effects degraded the effectiveness of the second and subsequent RVs, then one referred to "fratricide."

The issue was important at the time because estimates were that the probability of destruction was high (e.g., 90 percent) if the fratricide effects were small. Figure J.1 shows the difference between the independent-event calculations and a calculation accounting for fratricide. It assumes a reliability of 0.85, a single-shot kill probability of 0.65, and 1,000 silos. The expected number of surviving ICBMs more than doubles if fratricide is strong.

This calculation treats the probability of fratricide as a mere parameter, but where did the value come from? A number of scientists attempted estimates of the fratricide effect, but the most insightful work was probably that accomplished with detailed simulations that considered flight dynamics, shock waves, the nature of the dust cloud, the characteristics of the RVs themselves (ruggedness, ballistic coefficient, and so on), and the spatial configuration over time of both targets and attacking RVs. Even first-rate minds, when attempting to understand the problem analytically and physically, failed to account for what turned out to be major effects.

One lesson from this work was the value of combining both analytical work (relatively simple models of the phenomenon designed primarily to structure the issues) and detailed simulation. Another lesson was that the mathematics of the calculation mattered.

Saturation Effects in Ballistic Missile Defense

Scientists involved with ballistic-missile defense have recognized for decades that the Achilles' heel of many defenses is the ability of the attacker to

saturate them. Even if a defense battery is reliable and effective (big ifs, to be sure), it has a limited capacity for handling multiple targets and a limited inventory of surface-to-air missiles. It can be overwhelmed.

This has been known to be important in ballistic-missile defense calculations for at least three decades. Indeed, it is the principal reason that scientists have generally been pessimistic about the feasibility of an adequately effective one-layer defense. For example, if one tried to defend an aircraft carrier against nuclear-armed ballistic or cruise missiles, the effort would be hopeless against an opponent able to defeat the local defenses, something that typically requires many fewer weapons that one might think because of leakage.

There are many other "configural effects" in BMD. For example, if early warning radars can "warn" smaller-area radars of incoming weapons, then the defenses can be much more effective. But the early warning radars are then a critical node: that is, there are strong dependencies between terminal-defense capability and the continued existence of the early warning radars.

Gaming Effects

Yet another "configural effect" in ballistic missile defense involves the game theory of tactics. Suppose that there are 10 targets to be defended with 10 perfect interceptors. And suppose there are 10 perfect attacking missiles. How should the defending missiles be allocated among targets? How should the attacking missiles be allocated? If the defender has one missile for each target, then the attacker can destroy 5 targets with confidence by merely double-covering half of the set. Knowing that, how should the defender defend? Variants of this can be a complex game theory problem with nonintuitive results a factor of two or so different from what one might naively expect (or infinitely different in the example, where the number of targets killed could be 5 instead of 0).

Concentration and Counter-concentration in Operational-level Ground Combat

Although corps- or theater-level ground combat may not seem to involve configural problems because of its much higher aggregation, it does. The attacker does not attack uniformly, and the defender, if able to do so, responds nonuniformly to events. If one tries to calculate attrition, movement, or even cruder measures such as "who wins in a given corps sector?", and if one tries to use overaggregated equations that feature the overall force ratio, the results are nonsense. The attacker will concentrate and create a large force ratio in his main corridors. If the defender is fast enough, he will "reequilibrate," but if the attacker is fast enough, he will break through before that happens and win the battle.

At lower levels (higher aggregation), the same phenomenon occurs, but the

time scales are different. It is often believed by army officers that "reequilibration" is feasible—on average—at, say, the battalion level. That is, if one battalion is broken through by a locally concentrated force, then brigade and division-level reserves are supposed to come in and stem the breach. That is what command and control is all about for ground forces. But it may not happen that way. While aggregate models like Vector, TACWAR, and JICM do not represent this explicitly, the defender might lose at a force ratio that "ought" to be sufficient because the defender is unable to counter-concentrate quickly enough. In entity-level simulations like Janus, this is easier to see perhaps, and if one introduces statistical distributions, then many interesting things become visible. On the other hand, if one understands this well enough, one can reflect it in the higher-level aggregate models by calibrating the equations so that "break-even" occurs at a smaller force ratio than the naive calculation would suggest. This is not enough, however, if one wants to pay attention to probabilities. And so on. All of this has been understood to some extent conceptually for a long time. However, many of the configural effects have not been built in except at the entity level, where it is hard to avoid doing so (Davis, 1995b). Even there, workers too often focus on expected-value results rather than examining the probability distribution of results, which may in fact be multimodal.

DISCUSSION

These examples demonstrate that the configural effects noted in mine warfare theory are not only real, but akin mathematically to effects long recognized as fundamental in other domains of defense analysis. It should not be controversial to observe that proper modeling of combat must account adequately for probabilistic dependencies.[3] In mine warfare (and also in subjects such as penetration of air defenses), this is typically not understood intuitively because, for example, the n-th ship going through a mine field may be independent of the previous $n - 1$ ships, thus suggesting validity of an independent-events calculation. The subtlety here is that the minefield through which the ships move does not change significantly from ship to ship. Thus, the penetration probabilities are correlated (i.e., there are probabilistic dependencies). The resulting probability distributions have been shown by Timothy Horrigan to be distinctly multimodal in many real-world cases, and not at all like what one might expect from a primitive treatment of probabilities. Working the problem correctly can have large effects on both weapon requirements (how many to buy) and operational assessments during war.

Unfortunately, there are numerous cases throughout defense analysis where

[3]These examples do not do full justice to the range of effects treated in Horrigan (1991) and elsewhere, but they may demonstrate the ubiquity of dependency issues.

probabilistic dependencies are not accounted for properly. There are many reasons, including simplicity and familiarity of naive independent-event calculations, which many people can perform adequately with a spreadsheet. A deeper reason is that untutored intuition is often poor on issues involving probabilistic calculations. Indeed, probabilistic effects are sufficiently nonintuitive that workers often revert to simplified and naive calculations (including deterministic calculations when they are clearly inappropriate) even though they were once sensitive to the subtleties.

A third reason for failure to treat dependencies is that past computing capabilities did not permit workers to handle them effectively. This, coupled with the intuitive plausibility of many correlated effects averaging out or at least greatly simplifying, led to dependence on aggregate expressions such as the Lanchester equations described elsewhere in this report. That simplification has sometimes been valid and sometimes not. The conclusion here should be that without detailed analysis, and in many cases detailed simulation akin to "experimentation" with a real system, we should be skeptical about the validity of formulations that do not treat statistical dependencies.

Where might we find instances of such problems? The answer is "in virtually every part of combat modeling." Indeed, whenever calculations of effects such as attrition are effectively multiplying together a set of planning factors, the effect is to assume an independence of events that may not be correct. Current-day assessment of precision-strike effectiveness against an invading army is a prime candidate for errors.

CONCLUSIONS

The panel was asked a number of questions about configural effects in mine warfare. It was surprised that there continues to be controversy about the importance of treating correlation effects such as those that are manifest in mine warfare. When two ships move through a mine field, the mines do not re-randomize their locations between ships. Nor do the ships move independently (e.g., one may follow the trail of another). Nor, in fact, is the pattern of mines in a given waterway random in many circumstances. And so on. The probabilistic effects reported for mine warfare are real, and Navy doctrine and decision aids for dealing with mine and countermine warfare should reflect them. This is without prejudice to how that is accomplished, since there are a variety of possible modeling approaches. In particular, one approach is analytical and has advantages for moving from requirements (e.g., a maximum loss rate to mines) to estimates on how many mines to buy or how to lay them. A second approach involves simulation, which has many advantages (and is in some respects easier), but which is not easily able to answer the questions analytical models are best for. Also, analytical models and the related theory can clarify the structure of the problem so as to illuminate the dependence on controllable variables.

K

M&S-related Education

Donald Gaver, Naval Postgraduate School

BACKGROUND

The panel recommends an increased effort to educate future officers (and civilians) for work involving modeling and simulation (M&S) (discussed in the text of the report). Some of the work will involve developing M&S. Some will involve applying it. The applications will be in acquisition, training, and operations—each of these broadly construed. In many cases, the applications will be "analytical" in some sense—e.g., investigating the potential value of a new weapon system or tradeoffs among platforms, constructing a training activity or exercise that will expose participants to the desired range of situations and stresses, or assessing alternative courses of action. In other cases, the work will be more developmental or technological (e.g., managing a program that includes a model-building component or managing the assembly of a distributed interactive simulation specifically tailored to an exercise). The purpose of this appendix is to discuss the subjects to which students might be exposed to prepare them for such activities.[1]

DISTINGUISHING AMONG CLASSES OF EXPERTISE

If we contemplate the range of military or civilian professionals who will be working extensively with M&S, it quickly becomes evident that there are some distinct specializations. One useful breakdown from an M&S-centered perspective (one partly motivated by the discussion in Chapter 6 about layered architecture for M&S) is as follows:

[1]This appendix has benefited from inputs by Bernard Zeigler and Paul Davis.

• *Applications-oriented professionals (modelers, analysts . . .)* working in particular application domains such as acquisition, training, or operations, who can effectively pull together M&S assets as needed for their problems. In some cases, they will build or substantially alter models themselves; in other cases they will use preexisting models. They may or may not (and usually will not) consider themselves M&S specialists, often preferring to be identified as analysts, but, at least in the future, will have considerable M&S expertise.

• *Program managers* in domains that require overseeing applications-oriented M&S development or model-supported analysis.

• *Hardware-centered professionals* trained in the technologies of computers, networks, and related communications.

• *Software-centered professionals* trained in developing the software that utilizes the hardware to support the applications. Here we have in mind professional-quality software, not the computer programs typically generated by analysts or subject-focused modelers. Such software is intended for broad use, not just that within the originating group.

• *M&S facilitation specialists* trained to draw on technology and databases for both development and application of M&S in relatively complex contexts such as distributed interactive simulation, or developments exploiting model and tool repositories. These specialists would also be experts in assisting the collaboration with other professionals using groupware technologies of all types.

• *M&S scientists,* responsible for researching the architecture of both local and distributed M&S infrastructures, continually assessing their capabilities relative to future needs. These might be concerned about *n*-th generation "high-level architectures," complex computer-security issues, and tool development.

The focus in this appendix is on the applications-oriented professionals and program managers, not because they are more important than the others, but because it is here that the Department of the Navy probably wants to focus its special M&S-related education that is keyed to young officers. In contrast, the Navy Department will probably go to civilian employees or contractors for specialized skills in hardware, software, and so on. There will probably be an adequate supply of people with such skills, people who will have attended colleges and universities throughout the nation.

EDUCATION FOR FUTURE M&S USERS

A Perspective to Guide "Requirements"

Modeling and simulation (M&S) is one tool for use in making effective military decisions. A helpful planning or decision-assisting model is one that captures the essential elements of a situation or problem domain and that can be manipulated to provide synthetic experience efficiently. That experience is then

used as one input to guide choices of assets, tactics, or policy. An effective training model (or simulation) also provides synthetic experience, but now aimed to allow an operator or team to achieve and maintain particular skills. Parenthetically, decision-assisting models can provide invaluable training for decision makers, especially if either historical or hypothetical situations are presented that illustrate the realistic effects of uncertainty in its many aspects on the decision-making environment, and consequently on subsequent decision-affected outcomes.

If the above is an acceptable if abbreviated overview of the M&S enterprise, then one can ask for the background, sensitivities, and expertise desirable in a well-prepared professional user of M&S. The first observation is that today all such properties are unlikely to be embodied in one individual. What follows is a suggested order of priority for the types of talent and experience needed when an M&S enterprise is to be pursued. This list helps to define the educational needs.

Designer-Architect-Problem Formulator

No very substantial project involving M&S should be initiated without articulating one or more specific issues or questions to be examined. These questions should relate to the purposes of the organization guided by the decision maker to be advised, and should be as focused as possible. It requires art and experience to identify such questions; skill comes with practice. This arena is the purview of the essential designer-architect-problem formulator, whose proposals and direction set the stage for subsequent more technical modeling steps.

There is scattered literature useful to educate such specialist-generalists; some classics like G. Polya's *How to Solve It* are useful to read, but "how to formulate it" is more to the point, and some intensive searching for and creation of useful teaching material are in order. There are a number of books and courses somewhat relevant to such matters, sometimes in an operations-research or policy-analysis curriculum, sometimes elsewhere. Experience suggests that the case study approach is particularly valuable, because one learns how to formulate and conduct studies more by doing than by merely hearing principles. Accessing historical examples is a natural way to proceed; mining and refining corporate memory in particular areas and organizations can be undertaken to record lasting "lessons learned." Students can also be tasked to conduct "quick-response studies," which can be effective in instilling recognition that much can be done quickly with a mix of brainstorming, simple or relatively simple models, and clear problem-focused thinking.

On a formal-training level an initial educational background of natural science such as physics, chemistry, and electrical engineering, but also applied mathematics and statistics, has often been useful, particularly if the individual "likes problems" and enjoys the uncovering and exploitation of obscure structure and mechanism. Courses in mathematical and applied probability modeling, if designed

around problems (i.e., rather than being designed to illustrate sophisticated mathematical techniques) can be effective. Project work and informed mentoring are key to training new practitioners. Delivery of formal education or training in this essential M&S function has not been well addressed in very many places.

Conceptual Modeling

Once the desired questions are framed, choice of formal representation(s) of the problem elements must be made. It is often good practice to maintain several alternatives; for example, an initial low-resolution but fast and agile model could be explored, subsequently selectively enhanced by a more detailed higher-resolution followup, or a deterministic approach could be followed by a stochastic version. It is also valuable for students to be exposed to different perspectives of "the same domain," perspectives such as entity-level simulation on the one hand and operations-research analytical models to guide resource allocation on the other. A goal here should be to teach students to recognize and appreciate the values of different perspectives and representations, rather than associating themselves emotionally and nearly exclusively with one or the other.

The student is traditionally made aware of a number of model-type tools, typically quantitative-mathematical in nature but more recently also visual and animated, particularly in the training arena. Parenthetically, there is complementary overlap: statistical use is being made of "data animation and visualization" for dramatically conveying messages buried in complex data structures. Such can also be done to expedite model exploration.

Classical Methods

Here are some traditional model types and modeling tools that M&S professionals should know and appreciate to varying degree (no single individual is likely to be deeply conversant with all):

- Mathematical programming and optimization; other search methods such as evolutionary programming or genetic algorithms;
- Probability models and stochastic processes; search theory; reliability models; queuing theory;
- Statistics: data acquisition, and data analysis;
- Spreadsheet languages; simulation languages, others;
- Monte Carlo methods;
- Decision and control theory and analysis;
- Artificial intelligence (AI); rule-based systems; knowledge-based simulation;
- Game theory, game-theoretic optimization in simulations, and adversarial knowledge-based models in simulations (especially under uncertainty);

- Dynamical systems; chaos and complexity ideas; cellular automata; and
- Human factors and performance, and human-computer interfaces.

Higher-level Issues of Design

An important subject that is *not* often taught well is model design, especially for complex systems. One recently developed subject that is quite relevant here is object-oriented modeling (as distinct from programming), including modeling of complex systems. Students can profit from a study of systems dynamics, a subject usually associated with MIT's Jay Forrester. A classic book that addresses the nature and modeling of complex systems is Herbert Simon's *Sciences of the Artificial* (Simon, 1996).

New subjects of considerable importance involve complex adaptive systems and agent-based modeling. There are some interesting popular and semipopular materials available, but no single best reference of which we are aware suitable for graduate education. Appendix B discusses many of the items of interest here.

Appendix E emphasizes the importance of multi-resolution modeling and the desirability of having model families, but there is very little in current curricula—and not much in the current literature—to prepare people for such work.

It will often be true that a real problem can be completely addressed or "solved" by employing some relatively simple classical model. This possibility should not be overlooked.

Model Choice and Adaptation

This is the stage at which computer-intensive tools are invoked and computer science ideas find a place. Some important topics in this domain include the following:

- Object-oriented programming and, more generally, software engineering;
- Computer architecture and operating principles;
- Computer and communication networks; security;
- M&S tools and practices to aid comprehensibility, traceability, and "explanation";
- M&S designs to encourage and facilitate "exploratory analysis" amidst great uncertainty;
- Virtual-world and simulation systems;
- Distributed (operating) systems; and
- Virtual reality, and distributed interactive systems (DIS).

The M&S practitioner may well be required to address problems with versions of existing, even Service-specific models, such as the Navy's ITEM (or NSS), Army VIC or EAGLE, or Air Force THUNDER, or higher-level joint

systems such as TACWAR and JICM, and JWARS and JSIMS when they become available. He must be able to adapt these to particular situations and questions, being critical concerning results obtained. This step also includes basic parameter specification.

Model-output Analysis: Analytical Advice to Decision Makers

Model-output analysis is an extension of the above that includes the planning of model runs so as to economically obtain the necessary overall picture of response possibilities. Some of the above should be analyst-induced or instigated, while some can be in dialogue style with decision makers.

In summary, the professional user of M&S is desirably, but not currently realistically, responsible for a broad spectrum of knowledge and skills. This requires intensive, specific, and well-designed educational input with deliberate breadth and focus on the true usefulness of various viewpoints and technical tools. The field is bound to grow, and competition for appropriate military analysts so trained will grow also, but the opportunities should attract high-quality students and prospective practitioners.

VENUES FOR EDUCATION

Most of our discussion here relates primarily to postgraduate education in universities (e.g., in master's or Ph.D. programs), but an increasingly important part of educational strategy for organizations such as the Department of the Navy is the part that makes available specialized courses on an as-needed basis—e.g., before an officer takes on an assignment overseeing M&S development, exercise design, or weapon system analysis. Many possibilities exist here, ranging from short courses to self-learning packages. They are not complete substitutes for traditional degree studies, but they can be powerful supplements and, in some cases, partial substitutes. Continuing education is becoming a business-as-usual aspect of life for many knowledge workers, whether in or out of uniform.

L

Acronyms and Abbreviations

ACTD	Advanced concept technology demonstration
ADS	Advanced distributed simulation
BFS	Battlefield spreadsheet
C^4ISR	Command, control, communications, computing, intelligence, surveillance, and reconnaissance
CAD	Computer-aided design
CAM	Computer-aided modeling
CFOR	Command forces
CINC	Commander-in-chief
CMMS	Common models of the mission space
CNO	Chief of Naval Operations
COEA	Cost and Operational Effectiveness Analysis
CONUS	Continental United States
DARPA	Defense Advanced Research Projects Agency
DDR&E	Director of Defense Research and Engineering
DIA	Defense Intelligence Agency
DIS	Distributed interactive simulation
DMSO	Defense Modeling and Simulation Office
DOD	Department of Defense
DRB	Division ready brigade
FLEETEX	Fleet training exercise
HLA	High-level architecture
IOC	Initial operational capability
JCOS	Joint Countermine Operational Simulation

JCTS	Joint Tactical Combat Training System
JICM	Joint Integrated Contingency Model
J-MASS	Joint Modeling and Simulation System
JMEM	Joint munitions effectiveness manual
JROC	Joint Requirements Oversight Council
JSIMS	Joint Simulation System
JTF	Joint task force
JWARS	Joint Warfare System
JWCA	Joint Warfare Capabilities Assessment
M&S	Modeling and simulation
MEU	Marine expeditionary unit
MIP	Mixed initiative planning
MOE	Measure of effectiveness
NRAC	Naval Research Advisory Committee
NRaD	Naval Research and Development Division
NSS	Naval Simulation System
OMT	Object model template
ONR	Office of Naval Research
OSD	Office of the Secretary of Defense
OT&E	Operational test and evaluation
PC	Personal computer
RESA	Research, evaluation, and systems analysis (system)
RISTA	Reconnaissance, intelligence, surveillance, targeting, and acquisition
RMA	Revolution in military affairs
ROI	Return on investment
RSAS	RAND Strategy Assessment System
RTI	Run time architecture
SAFOR	Semiautomated forces
SAM	Surface-to-air missile
SBA	Simulation-based acquisition
SBD	Simulation-based design
SSBN	Nuclear-powered ballistic missile submarine
STOW	Synthetic theater of war
TMD	Theater missile defense
TOR	Terms of reference
UAV	Unmanned aerial vehicle
VE	Virtual engineering
VV&A	Verification, validation, and accreditation
WARSIM	Warfighter's simulation
WMD	Weapons of mass destruction (nuclear, chemical, and biological)
WSSF	Weapons Software Support Facility